papers and proceedings

the 11th

FEDERAL FORECASTERS CONFERENCE

2000

sponsored by

Bureau of Economic Analysis
Bureau of Health Professions
Bureau of Labor Statistics
Bureau of Transportation Statistics
Department of Veterans Affairs
Economic Research Service
International Trade Administration
National Center for Education Statistics
U.S. Census Bureau
U.S. Geological Survey

Federal Forecasters Conference - 2000

Papers and Proceedings

Edited by
Debra E. Gerald
National Center for Education Statistics

U.S. Department of Education
Office of Educational Research and Improvement

Federal Forecasters Conference Organizing Committee

Stuart Bernstein
Bureau of Health Professions
U. S. Department of Health and Human Services

Howard N Fullerton, Jr.
Bureau of Labor Statistics
U.S. Department of Labor

Debra E. Gerald
National Center for Education Statistics
U.S. Department of Education

Karen S. Hamrick
Economic Research Service
U.S. Department of Agriculture

Elliot Levy
International Trade Administration
U.S. Department of Commerce

Stephen A. MacDonald
Economic Research Service
U.S. Department of Agriculture

Jeffrey Osmint
U.S. Geological Survey
U.S. Department of the Interior

Norman C. Saunders
Bureau of Labor Statistics
U.S. Department of Labor

Kathleen Sorensen
U.S. Department of Veterans Affairs

Ching-li Wang
U.S. Census Bureau
U.S. Department of Commerce

Clifford Woodruff III
Bureau of Economic Analysis
U.S. Department of Commerce

Peg Young
Bureau of Transportation Statistics
U.S. Department of Transportation

Federal Forecasters Conference Organizing Committee

(Front Row) **Peg Young,** Bureau of Transportation Statistics; **Elliot Levy,** International Trade Administration; **Debra E. Gerald,** National Center for Education Statistics; **Kathleen Sorensen,** Department of Veterans Affairs; **Stuart Bernstein,** Bureau of Health Professions; and **Karen S. Hamrick,** Economic Research Service. (Back Row) **Clifford Woodruff,** Bureau of Economic Analysis; **Howard N Fullerton, Jr.,** Bureau of Labor Statistics; **Jeffrey Osmint,** U.S. Geological Survey; **Stephen A. MacDonald,** Economic Research Service; and **Norman C. Saunders,** Bureau of Labor Statistics. (Not Pictured) **Ching-li Wang,** U.S. Census Bureau.

Foreword

In the tradition of past meetings of federal forecasters, the 11th Federal Forecasters Conference (FFC/2000) held on September 14, 2000, in Washington, DC, provided a forum where forecasters from different federal agencies and other organizations could meet and discuss various aspects of forecasting in the United States. The theme was "Forecasting, Policy, and the Internet."

One hundred and eighty forecasters attended the day-long conference. The program included opening remarks by Debra E. Gerald and welcoming remarks from Mike Pilot, Acting Associate Commissioner for Employment Projections, Bureau of Labor Statistics. Following the remarks, a panel presentation was given by Neilson C. Conklin, Director of Market & Trade Economics Division, Economic Research Service, U.S. Department of Agriculture; Signe I. Wetrogan, Assistant Division Chief, of the U.S. Census Bureau, U.S. Department of Commerce; and Andrew A. White, Director of the Committee on National Statistics, National Research Council. Stuart Bernstein of the Bureau of Health Professions presented awards from the 2000 Federal Forecasters Forecasting Contest. Jeffrey Osmint of the U.S. Geological Survey presented awards for Best Papers from FFC/99.

In the afternoon, nine concurrent sessions in two time slots were held featuring a panel and 28 papers presented by forecasters from the Federal Government, private sector, and academia. A variety of papers were presented dealing with topics related to agriculture, the economy, health, labor, population, and forecasting software. These papers are included in these proceedings. Another product of the FFC/2000 is the *Federal Forecasters Directory 2000.*

Acknowledgments

Many individuals contributed to the success of the 11th Federal Forecasters Conference (FFC/2000). First and foremost, without the support of the cosponsoring agencies and dedication of the Federal Forecasters Conference Organizing Committee, FFC/2000 would not have been possible. Debra E. Gerald of the National Center for Education Statistics (NCES) served as chairperson and developed conference materials. Peg Young of the Bureau of Transportation Statistics (BTS) and Stephen M. MacDonald of the Economic Research Service (ERS) were instrumental in securing panelists for the morning session. Debra E. Gerald prepared the announcement and call for papers and provided conference materials. Norman Saunders served as the photographer for the day-long conference. Stuart Bernstein of Bureau of Health Professions organized and conducted the Federal Forecasters Forecasting Contest. Stephen M. MacDonald (ERS) and Peg Young (BTS) provided conference materials. Karen S. Hamrick (ERS) served as program chair and organized the afternoon concurrent sessions. Ching-li Wang of the U.S. Census Bureau provided conference materials. Jeffrey Osmint of U.S. Geological Survey provided conference materials and produced special awards for the forecasting contest and best papers. Kathleen Sorensen of the U.S. Department of Veterans Affairs provided conference materials. Howard N Fullerton, Jr. of the Bureau of Labor Statistics secured conference facilities and handled logistics. Also, recognition goes to Clifford Woodruff of the Bureau of Economic Analysis for support of the Federal Forecasters Conference.

A special appreciation goes to Peg Young and the International Institute of Forecasters for their support of this year's conference.

A deep appreciation goes to Paul Campbell, Sr. of the U.S. Census Bureau for reviewing the papers presented at the Tenth Federal Forecasters Conference and selecting awards for the FFC/99 Best Conference Papers.

Many thanks go to Linda D. Felton and Tina Terry-Eley of the Economic Research Service, U.S. Department of Agriculture, and Dave Walton of the U.S. Department of Veterans Affairs for directing the organization of materials into conference packets and staffing the registration desk.

Last, special thanks go to all presenters, discussants, and attendees whose participation made FFC/2000 another successful conference.

2000
Federal Forecasters Conference
Forecasting Contest

WINNER

Mirko Novakovic
Bureau of Labor Statistics

HONORABLE MENTION

Patrick Walker, Administrative Office of the U. S. Courts
Paul Campbell, U.S. Census Bureau
Peggy Podolak, U.S. Department of Energy
Betty W. Su, Bureau of Labor Statistics
Art Andreassen, Bureau of Labor Statistics

1999 Best Conference Paper

WINNER

"Contingent Forecasting of the Proportion with Small Incomes in a Vulnerable Nonmetro Population"

John Angle
Economic Research Service

HONORABLE MENTION

"Fertility and Long-Term Economic Projections"

Timothy Dowd
Joint Committee on Taxation

R. M. Monaco
INFORUM/University of Maryland

Contents

MACROECONOMIC AND REGIONAL FORECASTING ISSUES

FORECASTING MODELS

FORECASTS, SOFTWARE, AND THE INTERNET

COMMODITY FORECASTS

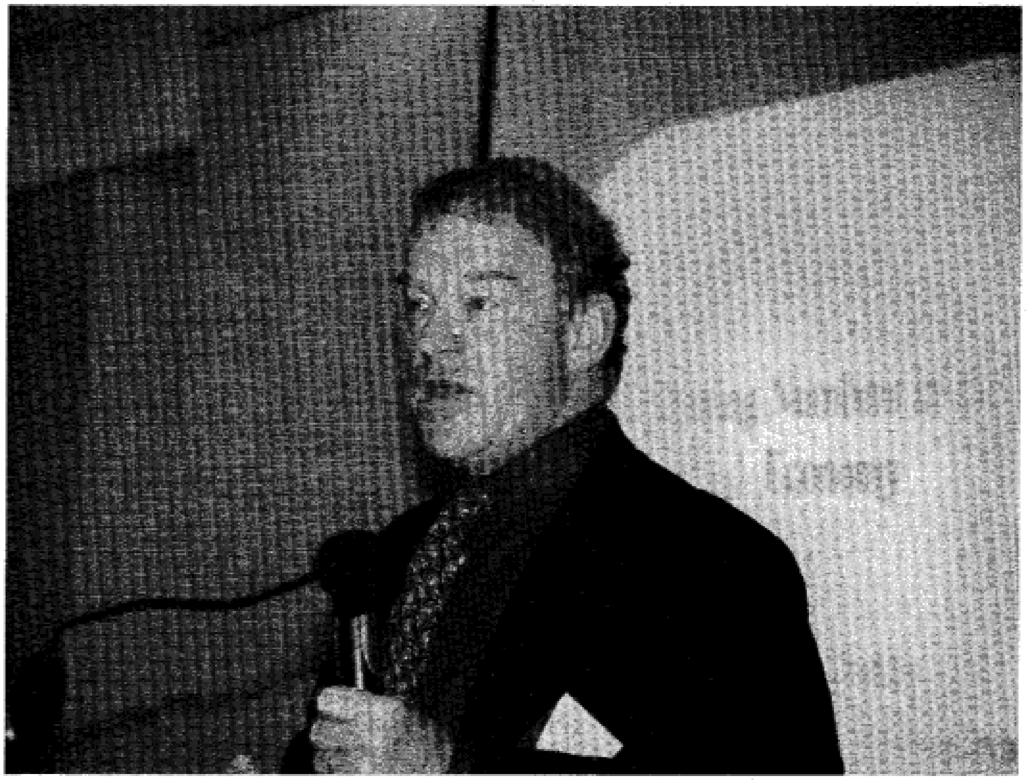
The 11th
Federal Forecasters Conference
FFC/2000

Scenes from the Conference

Photos by Norman C. Saunders, Bureau of Labor Statistics

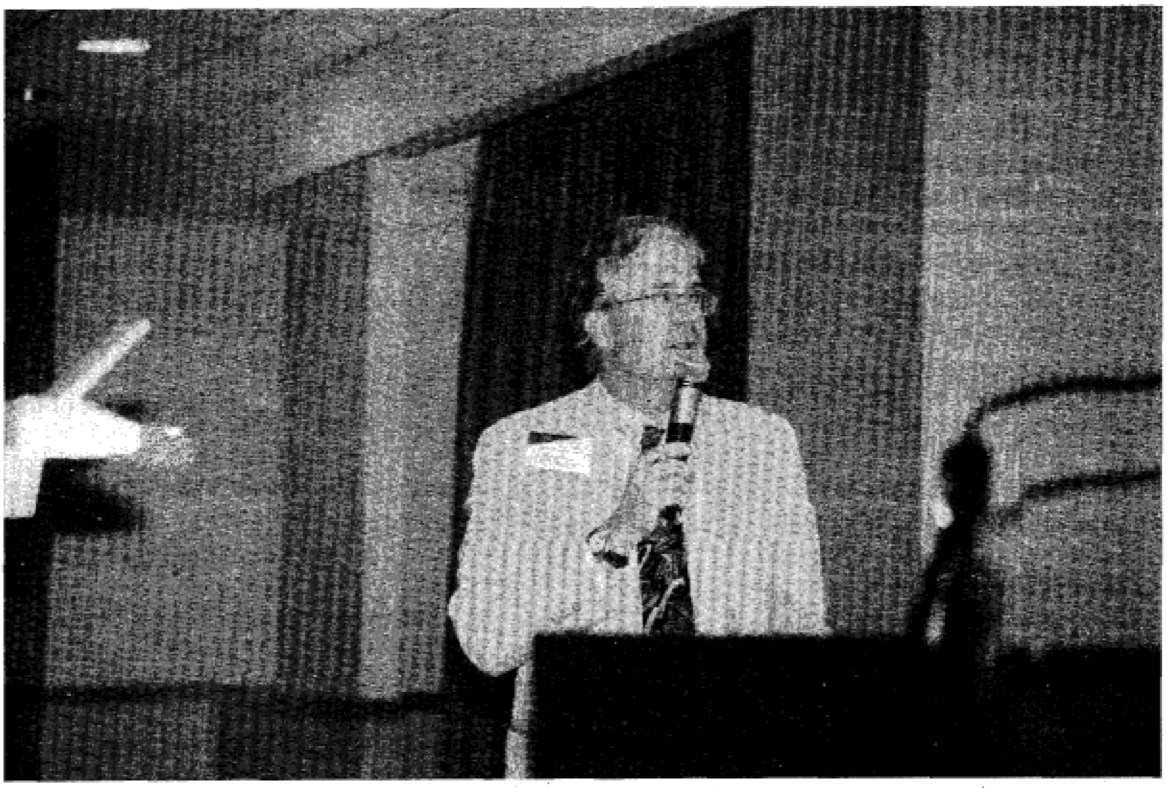
Mike Pilot of Bureau of Labor Statistics extends a warm welcome to the conference participants.

Peg Young of Bureau of Transportation Statistics introduces the morning panel.

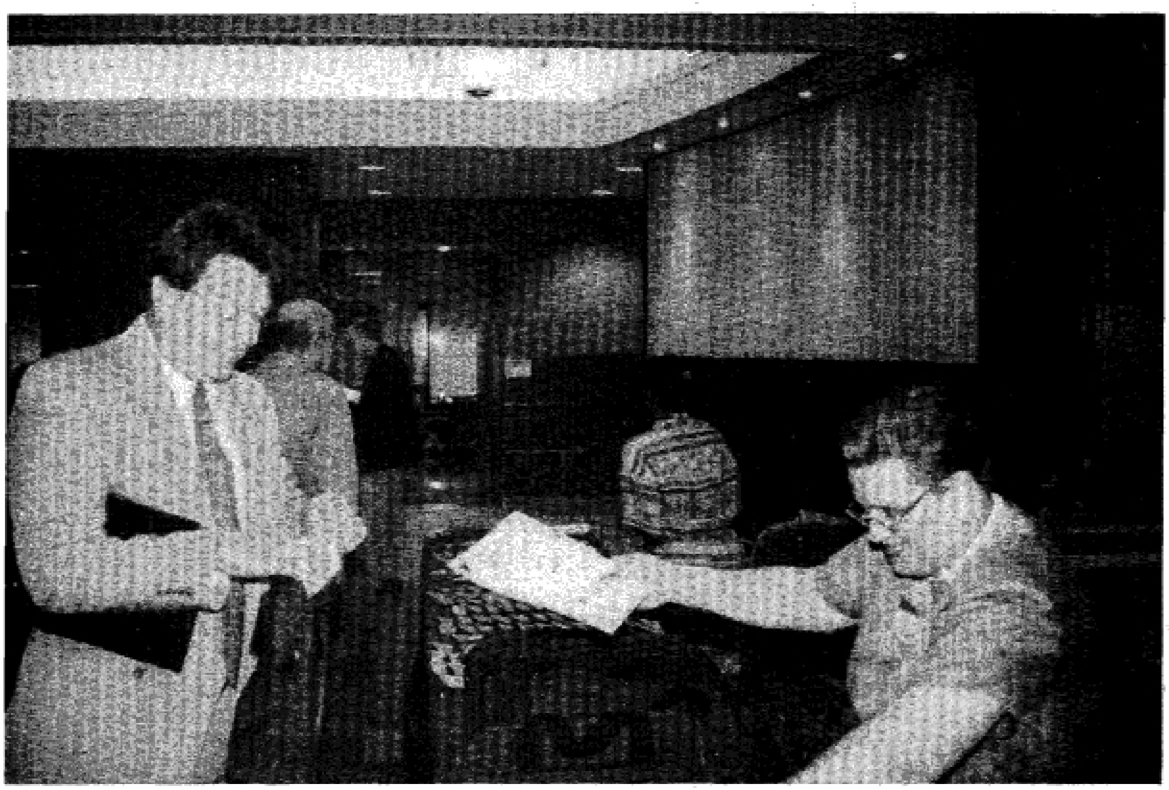


Neilson C. Conklin relates the ERS experience in delivering numbers to customers in the new economy.

Signe I. Wetrogan presents the opportunities and challenges facing the Census Bureau in disseminating data over the Internet.



Andrew A. White explores information technology research for federal statistics.



Linda D. Felton and Tina Terry-Eley greet participants and pass out conference materials.

Highlights of Panel Presentation

Digital Government and Federal Statistics

The Internet has greatly widened the base of potential customers for federal statistics and forecasts. Representing their respective agencies or organizations, the panel discussed the state of digital government with respect to its impact on federal statistics. Topics included the opportunities, as well as barriers, created by the electronic delivery of statistics.

Delivering Numbers in the New Economy

Neilson C. Conklin, Director, Market & Trade Economics Division
Economic Research Service
U.S. Department of Agriculture

Today there a new realities confronting statistical agencies in the "new" economy. Information is increasingly a public good; information technology is expanding the "reach" of private and public sector organizations, and allows us to deliver "richer" information to customers. Challenges are posed by the new economy and digital government. There is a breakdown in old delivery systems and relationships, which has already been experiences as a creative destruction in the private sector. Federal statistics agencies are undergoing new strategic thinking to find a new business model, in which we answer the questions:
Who are our customers?
How can we enrich our customers' experiences?
How do we design our products and services?
How do we allocate our resources?

The presentation related the ERS experience, as well as the issues for all of us in federal statistics, to show how to deal with transition issues (e.g., how do we meet our responsibilities to our 'unwired' customers), how to maintain quality control in a distributed environment, how to relegate the role of paper media in the future, and how to allocate resources.

Internet Use in Disseminating Population Estimates and Projections at the Census Bureau: Opportunities and Challenges

Signe I. Wetrogan, Assistant Division Chief for Population Estimates and Projections
Population Division
U.S. Census Bureau

It's no secret that access to and use of the Internet is continuing to increase. According to a Nielsen Media Research Survey taken in September 1997, 1 in 4 adults in the U.S. and Canada used the Internet – more than 58 million adults. The Census Bureau and, in particular, the Population Estimates and Projections Area take advantage of the Internet as a main mechanism in disseminating its data.

The Internet offers us the opportunity to quickly and easily release a large variety of data to a wide, multi-user group using various modes of delivery. At the same time, these opportunities pose many challenges, including the ability to reach this multi-user group and efficiently deliver the flexible types of data while preserving some type of data control, the ability to convey important data caveats and the need to archive revised datasets. This presentation outlined some of the steps that the Population Estimates and Projections Area is taking to meet some of these challenges.

Information Technology Research for Federal Statistics

Andrew A. White, Director
Committee on National Statistics
National Research Council

The National Research Council's Computer Science and Technology Board, in conjunction with the Committee on National Statistics, held a workshop on "Information Technology research for Federal Statistics" in early 1999. Participants in this workshop explored information technology (IT) research opportunities of relevance to the collection, analysis, and dissemination of federal statistics. The participants represented four broad communities: IT research, IT research management, federal statistics, and academic statistics. The workshop provided an opportunity for these communities to interact and to learn how they might collaborate more effectively in developing improved systems to support federal statistics. Highlights from the workshop summary report were discussed.

Concurrent Sessions I

POPULATION PROJECTIONS: CURRENT DEVELOPMENTS AND ISSUES

Chair: Thomas Bryan
U.S. Census Bureau

Discussant:
Peter D. Johnson
U.S. Census Bureau

U.S. Population Projections to the Year 2100
Frederick W. Hollmann, U.S. Census Bureau, U.S. Department of Commerce

Accuracy of the U.S. Census Bureau National Population Projections and Their Respective Components of Change,
Tammany J. Mulder, U.S. Census Bureau, U.S. Department of Commerce

Evaluation and Optimization of Population Projections Using Loss Functions,
Charles D. Coleman, U.S. Census Bureau, U.S. Department of Commerce

Projections of the Number of Households and Families in the United States: 1999 to 2025,
Ching-li Wang, U.S. Census Bureau, U.S. Department of Commerce

U.S. Population Projections to the Year 2100

Frederick W. Hollmann
U.S. Census Bureau

In January of this year, the Census Bureau released population projections for the United States from 1999 to 2100. While these projections yielded few surprises regarding the size and structure of the forecast population in comparison to previous series, the scope of the product was unprecedented. The most "eye-catching" change was the forecast horizon: for the first time, the projections reached as far as the end of the new century, to 2100. No previous series had ventured past 2080. We also expanded the level of demographic detail to include nativity, defined dichotomously as native versus foreign-born. Within each of these two categories, we produced the level of demographic detail, single year of age, by sex, by race, by Hispanic origin, that we have produced in the past. Last but not least, we increased the temporal density of the projections from annual to quarterly reference dates, primarily to allow users to select reference dates other than July 1. Finally, as in previous releases, we computed a "highest" and "lowest" variant of the projection series, based on extreme assumptions of all three of the major components of change. To a greater extent than in previous projection efforts, these series were intended to reflect the degree of uncertainty in the various components; hence, a relatively larger range was imposed on the relatively unpredictable migration component.

Far more interesting from the *producer's* perspective were changes in the methodology and assumptions underlying the projections. These are described in detail in a public document (Hollmann, Mulder, and Kallan, 1999). Throughout this paper, I will discuss projection assumptions that were described in this report, many attributable to the work of my two co-authors. Briefly, they are as follows.

1) We abandoned the assumption that international migration was constant over time with unchanging demographic composition. Instead, we viewed international migration in terms of the present distribution of migration by country (or country group) of origin, considering likely future developments from the major sources.

2) The addition of nativity as a differentiating variable in the projections allowed us to project emigration of the foreign-born through a schedule of rates, rather than as a constant matrix.

3) We adopted a target-based, rather than an extrapolation-based methodology for projecting mortality. While this did not result in a large change in the assumed levels, it addressed many of the technical problems present in earlier models resulting from quasi-independent projections of age-sex-race categories.

4) We reinstated the assumption that fertility rates would converge by race and Hispanic origin, abandoned in our penultimate release. However, unlike previous models that assumed convergence, we did not rest the convergence on present levels for the White, non-Hispanic population. Rather, we allowed all race and Hispanic origin cross-categories to trend toward a common target.

Migration to the United States

International migration to the United States is generally respected by demographers as the most difficult component to project, which is the reason that the "indefinitely constant" assumption is so frequently made. The most intimidating aspect of this component is most likely its dependence on policy, as well as its historic volatility. This volatility often results from events as unpredictable as foreign social and political upheavals. In recent years, we have witnessed millions of immigrants from Southeast Asia to the U.S., initially a result of the end of the war in Vietnam, later a result of the mass exodus of "boat people" to refugee camps in Thailand and elsewhere. From Cuba, we saw a boatlift of more than 100,000 "Marielitos" who arrived in the U.S. in 1980, primarily as a result of a policy shift from the Castro regime in Cuba. Our liberal policy regarding the admission of Cuban refugees and parolees, as well as refugees from the Soviet Union and its satellites has resulted in an ebb and flow of migrants from Cuba over the years since then–even as the Soviet hegemony in Europe (and the Soviet Union itself) disintegrated. Underlying this has been a steady stream of migration, legal and illegal, from Mexico and other portions of Central and South America of people seeking the relatively favorable demand for agricultural and other employment in the United States. In 1986, the Immigration Reform and Control Act (IRCA) effected the legalization of the residency of a large class of undocumented residents, clearing the way for them to become legal permanent residents, and ultimately U.S. citizens. With this improved immigration status came the right to sponsor other immigrants, primarily immediate relatives through family reunification, as well as new relatives through marriage. The Immigration Act of 1990 further promoted this process by exempting immediate relatives of U.S. citizens from numerical limitations. These factors resulted in a substantial increase in legal immigration during the 1990s. From the standpoint of a demographer attempting to project migration, this trend yields an interpretive problem: is secondary migration related to IRCA a self-feeding process that will continue to multiply the number of legal immigrants, or is it a historical event that will "run its course"?

The future of international migration was projected primarily on the following assumptions (Hollmann, Mulder, and Kallan, 1999).

1) The rapid increase in migration during the 1990s, driven in large part by the migration of relatives and most affecting the flows from Mexico and Central America, is transitory. Moreover, trends in economic development and reducing fertility in Latin America suggest this area will decline as a source of migration to the U.S. However, the presence of a Latin immigrant community in the U.S. will ensure its continued significant role. Undocumented migration across the Southwest border is characterized by an large excess of demand over "supply" (inability to prevent illegal entries), so we do not assume any future change, even if demand lessens.

2) Refugee movements from Southeast Asia, Cuba, and the Soviet Union will continue to decline in importance as a source of migration to the U.S., as our relations with these countries stabilize. Newer flows, principally from the former Yugoslavia and Africa, will see transitory "spikes". In the longer term (through the coming decade and beyond), refugee movements will decline.

3) Legal migration from some erstwhile less conventional sources will increase–especially in the long run. These include South Asia and sub-Saharan Africa. These areas are characterized by considerable potential for population growth, and (especially in the case of Africa) political instability.

4) Immigration policy affecting numerical limitations for employment-based visas will remain unchanged until 2020. After 2020, some increase in employment-based immigration is likely on account of the retiring baby boomers.

5) After 2030, we assume the overall level (but not necessarily the composition) of migration to the United States remains constant.

6) We reflected the uncertainty of all of these assumptions by projecting "low" and "high" series ranging from 0.58 million to 3.625 million per year in 2100, around a middle projection of 1.45 million. In 2020, the range is narrower: 0.56 million to 2.13 million around a middle projection of 1.09 million.

7) The age and sex composition of migration to the U.S. follows that of recent in-migrants by country of birth category. The race and Hispanic origin composition follows the composition by race and Hispanic origin of foreign-born migrants in the 1990 census by country of birth.

Emigration

For the emigration of foreign-born legal residents, we assume a schedule of rates by age, sex, and a very few region-of-origin categories sufficient to produce an average of 195,000 emigrants per year during the 1980s, based on research by Ahmed and Robinson (Ahmed and Robinson, 1994). This results in an annual emigration of 339,000 per year by 2020, increasing to 524,000 per year by 2100 (Hollmann, Mulder, and Kallan, 1999).

The use of rates to project emigration broke a conundrum that has haunted previous projection models. In projecting "high" and "low" values of emigration, should the numbers of emigrants be higher for a "high" net migration assumption, or lower? Considerations of forecast uncertainty suggest that the latter, since lower emigration supports higher net migration to the U.S. However, "demographic scenario" considerations favor the reverse. Emigration is largely a result of return migration of the foreign-born to countries of origin, hence, higher levels of foreign-born in-migration should be identified with higher levels of emigration, as more people are at risk of emigrating. When the emigration assumption is based on rates, this issue largely disappears. Clearly, a higher-growth model should feature lower *rates* of emigration for the foreign-born population. The foreign-born population can then function as a determinant of *numerical level.* In the present case, emigration increases fastest in the lowest series and slowest in the highest series; numerical levels converge in the late 2050s, then the trends proceed in opposite directions with emigration increasing fastest for the highest series. In the long run, the effects of the higher rate assumption in the lower series are overcome by the higher growth of the foreign-born in the higher series.

Emigration of U.S. natives is maintained as a constant distribution summing to 48,000 per year, based on research done by Edward Fernandez for the period around 1980 (Fernandez, 1995). Unfortunately, we have determined no credible way of trending this small component in the future.

Fertility

Past attempts by the U.S. Census Bureau to project long-term trends in fertility have engendered a well-studied skepticism among researchers (for example, Lee, 1999) who have pointed out that actual values of fertility often depart from between "high" and "low" forecast limits rather quickly. The trend over the past century has been one of fertility decline in the long run, coupled with enormous fluctuations of several years to a few decades in duration, settling to a comparatively constant trend in recent years. Recent values of the total fertility rate(TFR) have been close to the "replacement level" of 2.1, but slightly below it. We can isolate some major facets of our assumptions that are most critical to an understanding of our projections (Hollmann, Mulder, and Kallan, 1999).

1) We have tacitly rejected the notion that fertility in the U.S. will ape the recent history of Western Europe (most notably Italy and Spain), where the total fertility rate has reached levels only slightly above unity. It is our view that the European trends have been linked to increasing expectations of women for participation in economic life, as well as increasing rates of marital dissolution. In both of these areas, the United States has seen similar changes in the past, which may have been partially implicated in the fertility decline of the 1970s in the U.S., and (if at a different level) may explain some more recent declines in the more developed countries of Latin America. This history provides no basis for assuming a new American response to an admittedly pervasive phenomenon in 1990s Europe.

2) We assume convergence of fertility rates by race and Hispanic origin over time. This is partially justified by recent trends by race, specifically a long-awaited drop in the phenomenon of teen-age childbearing that has disproportionately affected the African American population. While evidence of a relative decline in Hispanic fertility has been absent, a projected decrease in the foreign-born component of the Hispanic population renders it nearly inevitable. Contrary to past series where convergence of fertility by race and Hispanic origin was assumed, we do not define the non-Hispanic White category as the target of convergence. Instead, we assume all race-origin groups to converge toward a target TFR of 2.1 by 2150 (50 years past the projection horizon).

3) While assuming convergence of fertility by race and Hispanic origin, we have held on to the assumption, also present in previous projection series, that race and Hispanic origin are principal determinants of fertility level. We assume this to the point of allowing our overall fertility assumption to be influenced by compositional effects. Otherwise stated, higher fertility race-origin categories will produce relatively larger generations of future mothers, so that overall fertility, will tend to rise more (or decline less) than would be implied by the assumptions made for individual groups. Specific to the present model, women of childbearing age will have an increasing proportion of Hispanic origin. Even after the effects of convergence, Hispanic women have somewhat higher rates of childbearing, so overall fertility increases as a result of this "bottom up" formulation of fertility by race and origin.

4) For the near to middle term we follow birth expectation data from the National Survey of Family Growth Cycle V(National Center for Health Statistics, 1995), but adjusted for the effect of future marital disruption and unfulfilled expectations (van Hoorn and Keilman, 1997). Taking account of the effects of changing racial and ethnic composition, this yields a TFR of 2.2 by the year 2025.

5) We assumed no differential fertility by nativity, meaning that we tacitly assumed that any nativity differentials were captured by the differential by race.

6) The range between the lowest and highest models reaches 1.9 to 2.6 by 2025, 1.6 to 2.7 by 2100, around a middle level of 2.2 for both years. These ranges reflect not only assumptions by race and Hispanic origin, but the compositional effects of changing demographic characteristics of women under the two extreme assumptions.

Mortality

Mortality is generally the component of population projections that requires the least speculative input. There is relatively little disagreement on the question of whether mortality in the U.S. should rise or decline: most researchers see it as declining. It is generally assumed that the focus of relevant public policy on longevity is to extend, rather than reduce it. Optimization of life expectancy is also generally the goal of the individual, although not necessarily the highest priority. (If it were, smoking would have ceased altogether, and people would not indulge in stressful, sedentary occupations such as estimating and projecting the U.S. population.) The epidemiological literature provides considerable explanation regarding the structure of mortality decline, defined by the trends in various causes of death, seen prospectively as risk factors.

The approach used to project the population of the United States in the present series was primarily actuarial, rather than epidemiological. We focused on the trend in the level and pattern of age-specific death rates by race and Hispanic origin, and the resulting trend in life expectation at different ages. This approach does not presume that trends in mortality from different causes are unimportant: rather it acknowledges that we are not able to forecast "turnarounds" in existing trends. Following are somewhat more specific characteristics of the mortality assumptions (Hollmann, Mulder, and Kallan, 1999).

1) We relied on a projection of life expectation at birth in the year 2065 by sex prepared by Lee and Tuljapurkar (Lee and Tuljapurkar, 1998), implying a level of 83 years for males and 88 years for females. This projection, in turn, was developed using the Lee-Carter methodology for projecting death rates based on historical trends (Lee and Carter, 1992). We assumed that similar rates of mortality decline continued from 2065 to the end of the century.

2) We derived an age pattern of mortality decline consistent with this life expectancy assumption based on expert opinion, from a survey of experts conducted by the Society of Actuaries in 1997 (Rosenberg and Luckner, 1998). These results indicated more rapid decline (relative to the base rate) for persons under 16 years of age than for persons age 65 and over, with the broad category of persons age 16 to 64 falling in between.

3) We assumed a constant-rate convergence of mortality by race and Hispanic origin from differentials that existed in 1997 to differentials of zero by the year 2150.

4) We assumed no differential mortality by nativity within cross categories of Hispanic origin and race.

Some Results, and the Public Reaction

Reaction to the new projections by the media tended to gravitate to two major observations that were featured in our press release. The first was the simple observation of the doubling of the population before the end of the projection horizon (by 2093). From the media perspective, it was easy to overlook the fact that the last century saw more than a tripling of the U.S. population, so we had projected a slowing of growth. The second was the increase in racial and ethnic diversity, specifically the emergence of Hispanics as the largest minority in the coming decade, and the acquisition of minority status by the non-Hispanic White population in the 2050s.

While the *critical* reaction to our projections was generally favorable, there were a couple of points in the assumptions that generated controversy. One criticism was that our fertility projections were too high–especially the presence of a gradual rise in fertility among non-Hispanic White women. This criticism was bolstered by the emergence of projections by the United Nations–released between the completion and release of our projections (United Nations, 1999). These indicated a convergence of fertility in industrialized countries toward levels significantly below replacement (although higher than those currently observed in Europe) by 2050.

A second class of critical reactions related to the migration assumption, and they were quite varied. One view held that international migration should increase indefinitely in proportion to the population; another held that it would be restricted by supply constraints, and would decline. Some criticism was explicitly directed to the concern that these projections would tend to fuel anti-immigration sentiment because of the juxtaposition of a modest migration assumption with apparent high population growth.

A very robust conclusion that, while not surprising, carried with it more social and demographic interest (in the opinion of this author) had to do with the foreign-born population. The proportion foreign-born increases gradually throughout the projection horizon, as one would expect. Far more interesting, however, is the relationship of the trend in the nativity of the population to the trend in race and Hispanic origin. We observe that while the non-Hispanic White and Black populations showed an increasing proportion foreign-born, the currently most heavily foreign-born race-ethnic categories--the non-Hispanic Asian and Hispanic populations–both show substantial declines in proportion foreign-born. This is a demographically robust finding, and arises from the fact that these immigrant-laden categories have age distributions highly favorable to childbearing within the United States, so that their second and higher-order generations of U.S. residents will make up ever increasing proportions of their numbers. This provides a clue to a much more important finding that can only be implied in the most qualitative terms. As racial and ethnic diversity increases, the way it is viewed by social scientists and ordinary citizens is likely to undergo fundamental changes. There is plenty of historic precedent for this, as other immigrant groups–principally people of European origin--have "melted" into a culture that tends to be defined primarily as "American".

As the population of the United States grows and becomes increasingly diverse, population projections provide an ever evolving view of these changes. We fully expect many of the findings in this series to be written over by new findings in future series that may arise from new censuses, new developments in vital events and international migration. Of greater concern to us as forecasters is the fact that we will see changes that result from methodological developments. Included will be changes in the way we approach race under the new rule of "check all that apply", as well as advances in the way we transmit characteristics across generations in the presence of increasing proportions of racially mixed marriages. This is definitely an exciting era for demographic forecasting.

References

Ahmed, Bashir, and Robinson, J. Gregory, *Estimates of Emigration of the Foreign-Born Population: 1980-1990,* Population Division Working Paper No. 9, U.S. Census Bureau, 1994.

Fernandez, Edward, *Estimation of the Annual Emigration of U.S. Born People by Using Foreign Censuses and Selected Administrative Data: Circa 1980,* Population Division Working Paper No. 10, 1995.

Hollmann, F.W., Mulder, T.J., and Kallan, J.E., *Methodology and Assumptions for the Population Projections of the United States: 1999 to 2100,* Population Division Working Paper No. 38, U.S. Census Bureau, December, 1999.

Lee, R.D., and Carter, L., "Modeling and Forecasting U.S. Mortality," *Journal of the American Statistical Association,* vol. 87-419, pp. 659-675, 1992.

Lee, Ronald D., "Probabilistic Approaches to Population Forecasting,", *Population and Development Review,* supplement to vol. 24, 1999.

Lee, R.D., and Tuljapurkar, S., *Population Forecasting for Fiscal Planning: Issues and Innovations,* unpublished, 1999.

National Center for Health Statistics, *National Survey of Family Growth, Cycle V,* 1995.

Rosenberg, M., and Luckner, W., "Summary of Results of Survey of Seminar Attendees", *North American Actuarial Journal,* vol. 2-4, pp. 64-82, 1998.

United Nations, *World Population Prospects: The 1998 Revision, Volume III, Analytic Report,* New York, 1999.

Van Hoorn, W., and Keilman, N., *Birth Expectations and Their Use in Fertility Forecasting,* Eurostat Working Papers, E4/1997-4, 1997.

ACCURACY OF THE U.S. CENSUS BUREAU NATIONAL POPULATION PROJECTIONS AND THEIR RESPECTIVE COMPONENTS OF CHANGE

Tammany J. Mulder
Population Division, U.S. Census Bureau

INTRODUCTION

Population projections are computations of future population size and characteristics based on separating the total population into its basic components of fertility, mortality, and migration and estimating the probable trends in each component under various assumptions (Srinivasan, 1998).[1] National forecasts give planners, legislators, policy makers, and researchers, among others, a glimpse of possible future demographic trends for the population and the forces acting to produce population change. Because forecasts are simply a compilation of reasonable assumptions as to what will happen to the current population in future years, the accuracy of forecasts will depend on the validity of the assumptions and the accuracy with which the assumptions are quantified. Correspondingly, it is critical for consumers of population forecasts to recognize the level of uncertainty found within population forecasts both in terms of their overall accuracy as well as in terms of the specific components of population change.

To date, the Census Bureau has not published a comprehensive analysis of the accuracy of their forecasts, which means customers depend on the expertise of the demographers producing the product. The aim of this research is to address this gap and systematically evaluate the accuracy of the existing Census Bureau forecasts both in terms of their ability to predict the national population as well as individual components of change.

The present paper evaluates the accuracy of Census Bureau population forecasts using an ex-post facto approach. That is, the performance of a forecast is evaluated relative to what was observed, which is operationalized here as intercensal estimates from 1947 to 1989, and the post-censal estimates from 1990 to 1999, produced by the Census Bureau (Byerly and Deardorff, 1995; Hollmann, 1990, 1993; U.S. Census Bureau, 1999, 2000a). In addition, the present study evaluates the assumptions used as input variables in the cohort component method. Specifically, this research will attempt to answer two research questions. First, how accurately did the Census Bureau forecast the total population and its respective components of change? Second, did the forecasts for the population and components produced by the Census Bureau perform more accurately than a naïve model assuming constant trends?

For the purposes of this research, the following terminology, which is consistent with language used among demographers and adapted from Smith and Sincich (1991), will be used to describe forecasts throughout the text:

Base year: the most recent estimate used to begin the forecast;
Target year: the designated point[2] (year) the forecast reaches;
Forecast period: the interval between first forecast year after the base year and target year;
Forecast error: the difference between the observed and the forecast population at a designated point in forecast period.

RESEARCH DESIGN AND METHODS

Choosing Among Multiple Forecast Series

In the recent past, the Census Bureau produced a middle series forecast and several alternate series based on differing assumptions for the components of change (fertility, mortality, and net immigration). Because the Census Bureau refers to the middle series as the "preferred series," and consumers commonly use this series, it is used hereafter for analytic purposes (U.S. Census Bureau, 2000b). For ease of discussion, each series will be identified by its respective *base year*. To evaluate the accuracy of the forecasts for the total population, seventeen forecasts were analyzed with base years ranging from 1947 to 1994 (U.S. Census Bureau, 1949 to 1996). Twelve series for the components of change are available from 1963 to 1994 (U.S. Census Bureau, 1964 to 1996).

Identification of a single middle series permits the comparison of error across products and the error experienced by each individual series. Therefore, in addition to analyzing the forecast error for each series,

[1] When discussing population projections, demographers often specify the difference between a "forecast" and a "projection." A projection generally represents possible population trends, while forecasts are produced to represent real population trends. In order to analyze the accuracy of the projections, we use the "preferred" middle series (U.S. Census Bureau, 2000b). In other words, this is the series the Bureau feels is most likely to take place, typifying a forecast. Furthermore, the object here is to analyze "forecast error," meaning the difference between forecast results and estimates. Therefore, the term forecast is used throughout the text.

[2] Throughout the text, "point" refers to a finite time interval within the forecast period.

the error for the combination of series at specific points in the forecast period are also calculated.

Error for the total population is measured for its annual percentage rate of change, or annual growth rate, which is calculated using the exponential formula.

Ex-post facto evaluation compares the forecast results with the historical population or component of change that was actually observed. Therefore, to evaluate the performance of past forecasts, each series is compared with intercensal (1947 to 1989) or postcensal (1990 to 1999) national estimates for the total population from 1947 to 1999. Both the estimated and the forecast population growth rates are calculated for annual intervals ending on June 30, while the components of change are summed for calendar years. Because few series forecast beyond 20 years in length, this analysis does not extend past the 20-year period.

Measurement of Forecast Error at Multiple Levels

A complicating factor in evaluating forecast error is that it can be calculated at different levels. It is possible to analyze an *individual point* in the forecast, the *individual series* to determine the error for specific products, as well as the error for *multiple forecast series* to assess the aggregation of error generally associated with the Census Bureau forecasts. In each case, *forecast error terms* -- the difference between the observed and the forecast population -- are used.

Forecast Error Patterns

Stated above, accuracy evaluation can be approached from two perspectives. Until now, the focus has been on evaluating *overall* forecast error. These evaluations relate strictly to the general performance of the forecast(s). The second, and more specific approach in performing a comprehensive assessment of forecast accuracy is that in addition to overall series error, there may also be patterns of error across time. In other words, how well did the forecasts perform throughout the length of forecast period and does a particular pattern exist? In order to assess the patterns of error throughout the forecast period, a supplemental analysis is presented for both individual and multiple series. Hereafter, *duration-specific forecast error* references the observation of patterns of error. Indicators used to measure overall error also measure the duration-specific forecast error for both the individual and multiple series.

Explanation of Indicators

Statistics used to measure the accuracy of forecasting methodology and assumptions originated from the analysis of economic forecasting. Demographers and statisticians apply these statistics to measure the accuracy of population forecasts at the national and sub-national level. Nevertheless, researchers have not reached a consensus as to which indicators are most indicative of the accuracy of national population forecasts (Ahlburg, 1992; Armstrong and Collopy, 1992). For the purposes of this analysis, the percent error (PE), the mean percent error (MPE), the mean absolute percent error (MAPE), the median absolute percent error (MdAPE), and the root mean squared error (RMSE) are used to measure accuracy.

These evaluative statistics apply to the individual and the multiple series analysis for both the overall forecast error and the duration-specific forecast error. To measure overall error, the PE is used to measure the forecast error that occurred at specified points in the forecast period (1, 5, 10, 15, 20 years). The MPE and the remainder of the statistics present the average within an individual series forecast period at specified intervals (5, 10, 15, and 20 year intervals). These indicators also measure the average across multiple series at designated points of the forecast period (1st, 5th, 15th, and 20th year from the base) as opposed to within series averages. Duration-specific forecast error is measured using the same indicators; however, for multiple series each indicator is analyzed annually (for each point) as opposed to designated points.

Comparison of the Census Bureau Forecast Models with a Naïve Model

Each Census Bureau forecast is based on a complex set of assumptions about how patterns of fertility, mortality and migration will behave over time. In order to understand the uncertainty related to these assumptions, each component of population change, as well as the population growth rate, is compared with a "naïve" model. Comparing the forecasts with a simplified naïve model assuming no change in future trends provides a benchmark to evaluate and compare the error experienced by the forecast model (Keyfitz, 1977: pg. 230). The naïve model is created by assuming the annual growth rate for the total population, the crude rates, and total number for the individual components remained constant as of the base year or "jump-off" population for the forecasts. For example, annual growth rates for the forecasts produced from 1967 to 1990 in P25-381 are compared with the constant annual growth rate for 1966, the designated population base of that forecast. The naïve model for number of deaths, however, cannot be simply held constant, *as this would not be representative of actual* trends. The naïve numbers of deaths are recalculated for each series based on the associated forecast population and the constant crude death rate.

RESULTS

Total Population Growth Rate Forecasts

Summary of Forecast Error for Growth Rates

Except for the 1974 and 1976 series, the pattern of under- and overestimation and level of accuracy for the individual series are closely related to the Census Bureau's assumptions for fertility and will be discussed in detail in the following sections. Tables 1 and 2 present the results of the indicators for each series. The first two forecast series, 1947 and 1949, greatly underestimated the growth rate as fertility rates began to rise in 1947, resulting in the Baby Boom. Short-term (five years) accuracy improved between 1953 and 1957 as growth rates remained at high levels resulting from high fertility rates. Following 1957, the growth rate began to decline, while the Census Bureau continued forecasting high growth rates. The total populations' forecast growth rates became more accurate within the recent past with average error statistics (excluding the MPE) falling below 10 percent within the first five years for the past five series as population growth stabilized in the 1980's and 1990's. The average error generally increased after the five year forecast period; however, the direction and magnitude of error did not increase or decrease in a consistent manner. Because of large outlier error terms, the multiple forecast error statistics do not represent the actual error experienced overall for the Census Bureau's forecasts.

In general, the naïve model outperformed the cohort component forecast, particularly in the latter half of the forecast period. Except for the 1957 series, the naïve model outperformed the forecast model for a minimum of one point in the measured forecast periods for each series. In contrast, recent cohort component forecasts consistently outperformed the naïve model in the first five years. The overall error remained high in comparison to a naïve model until the 1980's and 1990's.

Components of Population Change Forecasts[3]

Summary of Forecast Error for Fertility

The Census Bureau assumptions remained extremely optimistic about fertility trends remaining at levels experienced during the Baby Boom from 1963 to 1972, despite the continued decline experienced following the peak in 1957. As displayed in Table 3, error for total births decreased for series 1974 and 1976 because of two main factors. The 1974 series reduced the number of alternate series from four to three, resulting in one middle series with a lower completed fertility of 2.1, compared with an average of 2.5 and 2.1 for 1972. In addition, the number of births that actually occurred began to increase in the long-term forecast period. The 1976 series improved over the 1974 series by further reducing the short-term assumptions. In addition to a general improvement in the level of accuracy, the 1974 forecast began a trend of outperforming the naïve model of constant rates, with exception to the 1986 model.

In contrast, the 1982 and 1986 series were conservative and resulted in underestimating births. Series 1982 continued the use of the cohort fertility approach, while the 1986 series used a Box-Jenkins time series model for short-term forecasts. The completed fertility level was further reduced to 1.9 for 1982 and 1.8 for 1986. Following the 1990 turning point, the number of births remained stable. Accuracy improved for series 1991, which continued the use of the time series model, increased completed fertility to 2.1, and abandoned the racial convergence assumption, among other changes. This stability, combined with improved assumptions, permitted a more accurate forecast for those series produced within that decade. High levels of accuracy for short-term forecasts were duplicated for the 1992 and 1994 series, which abandoned the cohort method and assumed constant trends among the largest racial groups.[4]

The results of the comparison between forecast models differed for the number of births and the crude rate. The Census Bureau forecasts for the number of births were more accurate in the recent past. This is not necessarily true for the crude rate forecasts.

In summary, accuracy for the number of births improved in the recent past. Improved accuracy, however, does not seem to be explicitly determined by the different approaches toward deriving forecast assumptions (cohort vs. period) used to forecast short-term trends.

Summary of Forecast Error for Mortality

Beginning in 1963, the Census Bureau generally underestimated improvements in life expectancy. Error statistics for the forecasted number of deaths is presented in Table 4. Particular forecasts produced after 1976, in contrast, slightly overestimated improvement. Forecasts produced between 1963 and 1974 gradually increased in error, highlighting a trend of the Census Bureau's historically conservative approach toward forecasting improvements in life

[3] Error statistics for each component were calculated for both the total number and the crude rate. The results of the total number are presented in Tables 3 to 5. The results for the crude rates are not presented in this text.

[4] Fertility among non-Hispanic White, non-Hispanic Black, and non-Hispanic American Indian women remained at constant levels, while rates for Hispanic and Asian women were assumed to decline.

expectancy. Recent forecasts experienced superior performance in both overall and forecast period accuracy. This improvement in accuracy may be indicative of the stabilization of mortality trends in the late 1970's. In addition, the Census Bureau began producing a middle series mortality assumption; potentially further contributing to the overall level of mortality forecast accuracy. Similar to fertility, the error terms for the number of deaths are slightly larger throughout the forecast period than those for the crude rate as they are more dependent on the size of the forecast population. Multiple series forecast error generally increased throughout the forecast horizon, stabilizing after the 10th year of the forecast period. Lastly, except the three series, the naïve mortality models outperformed the Census Bureau forecasts. In comparison to fertility, the most recent forecasts, series 1992 and 1994, fail to exhibit superior performance relative to the naïve model.

Summary of Forecast Error for Net Immigration

Table 5 presents the error statistics for the forecasted number of net immigrants. Given that net immigration increased throughout the period between 1963 and 1999, the forecasts of constant rates were consistently underestimated. Error terms throughout the forecast period increased, and maintained the highest error statistics compared to the fertility and ~~mortality forecasts throughout.~~ Because most of the series begin with large forecast error terms within the first year, the base data used may be contributing to a large proportion of the error throughout the forecast period. Nonetheless, net immigration forecasts have improved in the recent past. This improvement is also evident when comparing the naïve and Census Bureau forecast models of net immigration. The naïve model consistently outperformed the Census Bureau forecast model, with exception to the fifth year average for 1991, 1992, and 1994, for both the number of net immigrants and the crude rate. In spite of this, the naïve results are not a dramatic improvement over the Census Bureau forecasts.

DISCUSSION OF RESULTS

This paper has evaluated the accuracy of population growth forecasts produced by the Census Bureau beginning with the 1947 series publication. To summarize the findings, the research questions asked previously are reiterated. First, how accurately did the Census Bureau forecast the total population and their respective components of change? In general, the forecasts produced by the Census Bureau overestimated total population growth. A detailed analysis of the components of population change, however, revealed a more complex pattern of over-and underestimation.

Erroneous assumptions about fertility following the Baby Boom era were largely responsible for a pattern of overestimation of the total population. Specifically, the growth rate forecast performance worsened for the series produced between 1957 and 1972. The number of births and the crude rate were severely overestimated between series 1963 and 1972, influencing the forecast growth rate. Before the 1957 series and following the 1972 series, annual growth rates were underestimated. Therefore, if the fertility component was not as grievously overestimated, the forecast results may be much more conservative and possibly underestimate the series as witnessed before the 1957 and after the 1972 series.

The mortality component of change generally presents the least amount of contributing error to the forecast model in comparison to fertility and possibly net immigration. The MAPE for both the number of deaths and the crude rates begin below 5 percent at the first year and never rise above 15 percent within the twenty year period.

The assumptions for constant levels of net immigration consistently produced underestimated series as the observed number of net immigrants continually increased for over thirty years. Forecasts were further troubled by the poor base data quality.

Recent forecasts for series 1991, 1992, and 1994, improve in accuracy over previous series within the first five years. Series 1991 and 1994's forecasts for fertility and mortality maintain smaller average error terms than previous forecasts, while the net immigration forecasts are smaller for the 1991 and 1992 series. This improvement in accuracy may be indicative of the stabilization of the components of change of the total population. In addition, the level of detail for the forecasts expanded as more race and Hispanic origin groups were added, the terminal age of the population data rose, and the quality of input data improved.[5]

The duration-specific forecast error generally increases throughout the forecast period for both multiple series and individual series for the growth rate and the components of change. The magnitude by which the error increased differs for each component of population change. Net immigration consistently

[5] Beginning with the 1991 series, the Census Bureau began producing forecasts with greater detail for race and Hispanic origin groups. The vital statistics data and the estimates were used to forecast four race groups by Hispanic and non-Hispanic origin (U.S. Census Bureau, 1993 (P25-1092)). In 1982, the age distribution of the forecast population was extended from 85 years and over to 100 years and over (P25-952). Lastly, for the 1991 series, the detail for net immigrants were expanded to five types of immigration to the U.S. (P25-1092).

maintains the highest level of error throughout the multiple series statistics, followed by fertility and mortality. Fertility increased rapidly within the first half of the forecasts; however, the stabilization of rates in the latter half is the result of an eventual increase in the fertility of American women, following a major decline. Mortality maintains the smallest error and remains stable throughout the forecast period past the tenth forecast year, as compared to the net immigration and fertility forecasts.

Secondly, did the forecasts for the population and the components of change produced by the Census Bureau perform more accurately than a naïve model assuming constant change? With exception to the recent forecasts of 1991, 1992, and 1994, and earlier series 1955, 1957, and 1963, the naïve models outperformed the Census Bureau forecasts for the growth rate and each component of population change. It is evident that the Census Bureau's inability to forecast turning points in trends greatly diminishes the accuracy of each forecast series.

The assumption of constancy for the naïve model outperformed the Census Bureau forecast assumptions for series experiencing a change in trends. In contrast, once the population stabilized in the recent past or experienced minimal to moderate change before the Baby Boom, the Census Bureau forecasts generally outperformed the naïve model.

CONCLUSION

Population forecasts produced by the Census Bureau are used widely, informing researchers, planners, legislators, and many others, on the future course of population change. Because forecasts are subject to inherent uncertainty, as they are based on a compilation of reasonable assumptions for the components of population change, it is essential to educate customers as to the amount of uncertainty within the forecasts for the population and the components of population change. Throughout the second half of the century, the forecasts produced by the Census Bureau improved in accuracy as a result of several factors including improvements in data quality and methodology. Nonetheless, this study reveals that forecasters failed to foresee turning points in population trends, resulting in erroneous forecasts, particularly for fertility and net immigration. In addition, with exception of net immigration, the assumptions formulated by the Bureau were often outperformed by simple assumptions of constancy.

Recent forecasts produced in the 1990's minimize the inherent uncertainty and provide a reliable product for consumers. The forecast reliability is, in all likelihood, the result of the stabilization of the components of population change.

In order to reduce uncertainty in future products, further analysis is necessary to understand the uncertainty in forecasting specific characteristics of the population, such as the forecasts of the race and Hispanic origin distribution and the age-specific assumptions for the components of change. Correspondingly, a detailed analysis comparing the specific assumptions made between products may strengthen the understanding of the weakness in the chosen assumptions.

REFERENCES

Ahlburg, Dennis A. 1982. "How Accurate are the U.S. Bureau of the Census Projections of Total Live Births?" *Journal of Forecasting* 1: 365-374.

_____. 1992. "A Commentary on Error Measures," *International Journal of Forecasting* 8: 99-100.

Armstrong, J. Scott and Fred Collopy. 1992. "Error Measures for Generalizing about Forecasting Methods: Empirical Comparisons," *International Journal of Forecasting* 8: 69-80.

Ascher, William. 1978. *Forecasting: An Appraisal for Policy-Makers and Planners*, Baltimore, MD: Johns Hopkins.

Byerly, E. and K. Deardorff. 1995. "National and State Population Estimates: 1990 to 1994," *Current Population Reports P25-1127* U.S. Government Printing Office: Washington, DC.

Hollmann, Frederick. 1990. *Population Estimates for the United States*. U.S. Census Bureau, Current Population Reports, P25-1045. Washington, DC: U.S. Government Printing Office.

Keyfitz, Nathan. 1977. *Applied Mathematical Demography*. New York, NY: Wiley and Sons.

Long, John F. 1987. "The Accuracy of Population Projection Methods at the U.S. Census Bureau," Presented at the Annual Meeting of the Population Association of America, Chicago, Illinois.

U.S. Census Bureau. 1949 to 1996. *Population Projections of the United States*. Current Population Reports, P25-18 to P25-1130. Washington, DC: U.S. Government Printing Office.

_____. 1999. "Population Estimates for the United States by Age, Sex, Race, Hispanic Origin, and Nativity: 1999 to 2100," Internet http://www.census.gov/ population/www/ estimates/uspop.html>.

_____. 2000a. *Unpublished Data from the National Estimates Product*.

_____. 2000b. "Population Projections for the United States by Age, Sex, Race, Hispanic Origin, and Nativity: 1999 to 2100," <http://www.census.gov/population/www/ projections/natproj.html>

Smith, Stanley and Terry Sincich. 1991. "An Empirical Analysis of the Effect of Length of Forecast Horizon on Population Forecast Errors," *Demography* 28(2): 261-274.

Srinivasan, K. 1998. *Basic Demographic Techniques and Applications*, Thousand Oaks, CA: Sage Publications, Inc.

Stoto, Michael. 1983. "The Accuracy of Population Projections," *Journal of the American Statistical Association* 78(381): 13-20.

Table 1. Error Statistics for the Forecasted Annual Growth Rate for the Total US Resident Population: 1947 to 1999.

[In percents. Resident population]

Forecast Periods	Individual Series (by Base Year)																	Multiple Series
	1947	1949	1953	1955	1957	1963	1966	1969	1970	1972	1974	1976	1982	1986	1991	1992	1994	
Five years																		
MPE (%)	(31.17)	(16.51)	(14.09)	(13.58)	0.52	12.93	12.91	(1.07)	13.37	(0.89)	(20.76)	(21.49)	3.88	(8.60)	1.93	7.62	(2.54)	(3.76)
MAPE (%)	31.17	18.52	14.09	13.58	1.98	14.06	12.91	14.16	20.10	4.08	20.76	21.49	3.88	9.88	2.51	7.62	3.30	15.04
MdAPE (%)	30.39	18.05	15.41	13.21	2.65	13.19	6.62	16.66	21.01	3.34	19.86	25.06	3.39	5.43	3.11	9.66	4.16	9.20
RMSE	0.57	0.37	0.28	0.24	0.03	0.18	0.16	0.17	0.22	0.05	0.21	0.25	0.04	0.14	0.03	0.08	0.04	0.30
RMSE Naïve	0.22	0.05	0.11	0.10	0.19	0.30	0.12	0.17	0.21	0.11	0.12	0.12	0.07	0.10	0.10	0.12	0.05	0.18
Ten years																		
MPE (%)		(30.81)	(15.62)	(9.83)	13.53	20.11	29.83	7.14	21.80	3.41	(9.33)	(8.66)	(5.05)	(17.53)				9.36
MAPE (%)		31.81	15.62	11.49	14.26	20.68	29.83	14.76	25.17	6.15	14.87	13.17	8.92	18.17				26.89
MdAPE (%)		39.33	15.41	12.18	5.61	16.26	26.89	15.38	27.24	3.22	17.61	7.84	4.98	23.69				23.66
RMSE		0.61	0.28	0.21	0.25	0.25	0.34	0.17	0.27	0.09	0.17	0.18	0.13	0.21				0.37
RMSE Naïve		0.05	0.11	0.25	0.42	0.33	0.16	0.13	0.19	0.10	0.12	0.10	0.08	0.10				0.30
Fifteen years																		
MPE (%)			(9.59)	2.54	26.56	31.40	35.61	15.33	31.93	10.93	(4.93)	(9.10)	(12.01)					23.94
MAPE (%)			12.88	16.76	27.05	31.78	35.61	20.41	34.17	12.76	11.88	12.50	14.59					34.91
MdAPE (%)			14.25	14.08	25.17	31.58	41.27	17.07	32.18	8.51	10.30	7.37	17.94					31.25
RMSE			0.24	0.24	0.39	0.37	0.40	0.23	0.36	0.15	0.14	0.17	0.18					0.39
RMSE Naïve			0.30	0.45	0.54	0.38	0.15	0.12	0.22	0.13	0.10	0.09	0.07					0.38
Twenty years																		
MPE (%)			(6.77)	9.91	41.60	37.00	44.53	20.62	32.23	7.54	(9.85)	(13.91)						23.44
MAPE (%)			12.21	20.57	41.96	37.28	44.53	24.43	33.92	12.46	15.06	16.46						37.78
MdAPE (%)			12.68	16.54	39.24	38.82	46.98	20.75	32.85	10.24	17.61	22.06						28.66
RMSE			0.22	0.27	0.54	0.42	0.47	0.26	0.35	0.15	0.18	0.20						0.43
RMSE Naïve			0.39	0.54	0.63	0.39	0.18	0.11	0.24	0.12	0.11	0.09						0.46

Source: Population Projections Program, Population Division, US Census Bureau: May 2000

Table 2. Percent Error for the Total U.S. National Population Forecasted Annual Growth Rates: 1947 to 1999

[In percents. Resident population]

Base Year	Percent Error (%) of Forecast Period				
	1st	5th	10th	15th	20th
1947	(12.69)	(48.62)			
1949	5.02	(35.80)	(47.27)		
1953	(6.23)	(15.41)	(14.25)	16.77	10.42
1955	(15.05)	(9.20)	8.30	14.08	37.13
1957	0.82	2.79	47.76	64.34	83.74
1963	(2.83)	29.16	50.66	46.20	69.88
1966	6.41	4.66	56.71	61.34	66.69
1969	(16.66)	20.30	10.99	47.59	27.44
1970	(16.83)	27.47	23.66	52.31	10.83
1972	(8.51)	3.09	20.72	21.04	(15.75)
1974	(26.49)	(18.09)	14.58	(5.08)	(26.08)
1976	(25.06)	(6.01)	2.23	(24.00)	(29.87)
1982	2.25	3.39	(24.95)	(31.25)	
1986	3.21	(22.31)	(27.44)		
1991	0.01	3.11			
1992	4.94	1.94			
1994	1.37	(4.35)			

Source: Population Projections Program, Population Division, US Census Bureau: May 2000

Table 3. Error Statistics for the Forecasted Number of Births for the Total US Resident Population: 1963 to 1999.

[Resident population]

Forecast Period	Individual Series (By Base Year)												Multiple Series
	1963	1966	1969	1970	1972	1974	1976	1982	1986	1991	1992	1994	
Five years													
MPE (%)	13.07	17.07	16.17	34.98	19.74	8.46	2.46	2.42	(8.34)	0.19	2.58	0.08	11.97
MAPE (%)	13.07	17.07	16.88	34.98	19.74	8.46	3.06	2.42	8.34	0.49	2.58	0.92	15.39
MdAPE (%)	14.77	16.17	19.26	39.24	21.55	8.11	2.37	2.62	10.24	0.50	2.59	0.95	9.42
RMSE	539,939	642,354	643,254	1,189,153	648,612	294,905	121,608	92,489	357,445	22,365	102,095	36,616	702,241
RMSE Naïve	465,722	84,814	352,048	513,132	100,102	184,225	337,241	78,290	261,370	162,579	147,085	49,386	346,913
Ten years													
MPE (%)	25.16	36.14	25.46	44.55	22.07	10.29	4.91	(1.47)	(9.32)				26.02
MAPE (%)	25.16	36.14	25.82	44.55	22.07	10.29	5.21	3.89	9.32				30.10
MdAPE (%)	22.36	34.55	31.53	51.18	23.57	11.06	5.93	3.04	10.11				23.39
RMSE	1,010,112	1,327,378	928,437	1,544,270	764,552	377,793	212,161	184,829	381,507				1,235,627
RMSE Naïve	603,648	338,150	356,068	453,133	235,180	367,932	447,414	278,243	204,582				495,133
Fifteen years													
MPE (%)	38.79	43.99	28.48	47.18	22.46	9.41	3.08	(3.46)					27.46
MAPE (%)	38.79	43.99	28.72	47.18	22.46	9.41	4.74	5.08					30.77
MdAPE (%)	32.85	55.81	33.42	52.18	23.07	10.62	4.86	6.88					28.19
RMSE	1,482,111	1,615,241	1,051,433	1,681,661	800,827	360,159	197,610	226,634					1,392,168
RMSE Naïve	724,413	301,234	291,243	372,748	333,905	505,295	610,719	268,718					577,276
Twenty years													
MPE (%)	44.27	47.80	29.38		19.23	6.39	1.15						24.51
MAPE (%)	44.27	47.80	29.56		19.23	7.73	4.71						27.07
MdAPE (%)	59.13	57.82	33.43		22.27	8.62	4.76						17.05
RMSE	1,687,995	1,777,148	1,100,525		722,383	316,578	194,288						1,370,479
RMSE Naïve	681,960	263,679	281,477		494,923	627,855	662,059						610,775

[1] The forecasted RMSE and Naïve RMSE are expressed as the number of births.

Source: Population Projections Program, Population Division, US Census Bureau: May 2000

Table 4. Error Statistics for the Forecasted Number of Deaths for the Total US Resident Population: 1963 to 1999.

[Resident population]

Forecast Period	Individual Series (By Base Year)												Multiple Series
	1963	1966	1969	1970	1972	1974	1976	1982	1986	1991	1992	1994	
Five years													
MPE (%)	2.35	3.40	4.01	7.60	7.85	10.51	5.43	(0.91)	0.75	(0.24)	(3.78)	1.17	4.51
MAPE (%)	2.35	3.40	4.01	7.60	7.85	10.51	5.43	0.91	1.25	0.93	3.78	1.29	5.05
MdAPE (%)	2.65	3.79	3.45	5.91	9.45	10.74	4.56	1.12	1.26	0.90	4.30	1.67	3.55
RMSE	47,814	73,116	85,993	155,663	164,774	202,725	107,745	21,869	29,102	23,844	90,270	34,328	128,743
RMSE Naïve	36,485	37,336	53,168	74,715	110,314	46,502	19,597	47,199	35,190	40,627	53,298	24,018	78,293
Ten years													
MPE (%)	3.57	6.38	8.55	11.07	10.46	10.96	6.40	(0.45)	(0.13)				9.20
MAPE (%)	3.57	6.38	8.55	11.07	10.46	10.96	6.40	0.96	1.13				8.73
MdAPE (%)	3.35	5.34	9.40	12.71	11.63	11.25	6.82	0.91	1.21				10.96
RMSE	75,907	144,696	187,922	227,281	215,685	216,213	131,063	24,133	27,557				200,461
RMSE Naïve	61,657	114,019	146,035	149,290	145,267	47,106	22,778	46,768	30,085				150,582
Fifteen years													
MPE (%)	6.78	9.27	10.35	12.51	11.14	10.72	7.11	(0.77)					11.36
MAPE (%)	6.78	9.27	10.35	12.51	11.14	10.72	7.11	1.11					10.97
MdAPE (%)	5.14	9.35	12.21	13.72	11.98	10.74	7.00	1.29					12.36
RMSE	159,959	206,115	222,970	258,269	231,206	216,818	151,764	28,141					241,556
RMSE Naïve	161,501	195,258	183,613	192,889	157,270	40,718	27,397	64,817					217,786
Twenty years													
MPE (%)	8.94	10.61	10.99		11.60	10.94	7.59						12.72
MAPE (%)	8.94	10.61	10.99		11.60	10.94	7.59						12.18
MdAPE (%)	7.70	13.38	12.78		12.11	10.89	7.55						13.15
RMSE	205,967	232,965	236,868		244,999	227,585	166,826						265,525
RMSE Naïve	233,990	252,523	205,619		172,701	41,704	42,064						278,889

Source: Population Projections Program, Population Division, US Census Bureau: May 2000

Table 5. Error Statistics for the Forecasted Number of Net Immigrants for the Total US Resident Population: 1963 to 1999.

Forecast Period	Individual Series (By Base Year)												Multiple Series
	1963	1966	1969	1970	1972	1974	1976	1982	1986	1991	1992	1994	
Five years													
MPE (%)	(22.23)	(10.07)	(7.14)	(7.47)	(9.03)	(22.41)	(35.22)	(28.52)	(21.59)	(1.58)	(1.04)	(8.38)	(20.79)
MAPE (%)	22.23	10.27	7.96	8.29	9.84	22.41	35.22	28.52	24.01	6.02	5.48	8.38	21.13
MdAPE (%)	24.62	11.70	2.04	2.04	6.76	23.81	35.06	30.61	17.84	6.09	6.09	4.67	19.35
RMSE	102,218	63,204	58,445	62,782	65,542	142,743	271,040	184,491	276,493	70,267	59,906	92,414	189,197
RMSE Naïve	54,944	41,180	49,723	91,459	64,866	149,788	245,622	49,605	210,064	91,180	128,113	100,299	145,237
Ten years													
MPE (%)	(27.33)	(8.09)	(14.77)	(17.82)	(23.91)	(30.40)	(35.13)	(33.69)	(31.24)				(36.53)
MAPE (%)	27.33	8.59	15.18	18.23	24.32	30.40	35.13	33.69	32.45				36.53
MdAPE (%)	26.91	5.07	14.25	14.25	27.66	32.28	33.92	31.98	38.53				35.06
RMSE	130,256	60,460	109,067	174,352	205,406	222,383	246,596	293,748	329,232				321,813
RMSE Naïve	78,158	48,212	78,830	132,608	204,651	229,378	219,918	183,427	222,725				244,045
Fifteen years													
MPE (%)	(30.30)	(17.13)	(22.65)	(23.73)	(28.10)	(33.68)	(37.28)	(38.92)					(44.64)
MAPE (%)	30.30	17.47	22.92	24.00	28.37	33.68	37.28	38.92					44.64
MdAPE (%)	30.07	13.61	23.81	31.51	32.45	35.06	38.32	36.80					42.91
RMSE	153,830	164,087	184,684	193,284	215,297	239,604	280,327	352,272					357,351
RMSE Naïve	101,711	134,951	148,298	127,309	214,459	246,936	254,173	230,613					304,553
Twenty years													
MPE (%)	(36.28)	(21.61)	(27.04)		(32.48)	(37.77)	(41.77)						(50.16)
MAPE (%)	36.28	21.86	27.25		32.69	37.77	41.77						50.16
MdAPE (%)	34.64	22.99	32.28		33.92	38.91	39.74						50.00
RMSE	231,952	179,534	209,551		279,627	313,470	349,784						423,619
RMSE Naïve	183,119	143,108	168,448		278,807	320,495	323,787						400,816

Source: Population Projections Program, Population Division, US Census Bureau: May 2000

EVALUATION AND OPTIMIZATION OF POPULATION PROJECTIONS USING LOSS FUNCTIONS

Charles D. Coleman
Population Division
U.S. Census Bureau
Washington, DC 20233-8800
Email: ccoleman@census.gov

1. Introduction

Loss functions are useful for the evaluation of population projections. (Coleman, 2000a and Bryan, 1999) They can incorporate trade-offs between numerical and percentage changes and compare areas of differing sizes on the same basis. This paper briefly discusses loss functions, then proceeds to the problem of developing point population projections which minimize the expected total loss of set of population projections for a given set of areas at a single point in time, provided that a subjective probability distribution function of the future populations can be constructed. These projections are based on Knightian risk, in that the probabilities are quantifiable. Knightian uncertainty enters into this problem, when there is residual uncertainty about the subjective probabilities or there exist events whose probabilities cannot be determined or whose possibilities may not even be known beforehand.

Section 2 briefly introduces the use of loss functions to measure the accuracy of cross-sectional projections. This Section begins by assuming the presence of an impartial decision-maker who has preferences over outcomes. Since this decision-maker is unlikely to exist, Webster's rule is proposed, as it possesses several desirable properties. (Coleman, 2000c)

Section 3 applies the techniques of Section 2 to finding the expected total loss associated with a particular set of cross-sectional population projections. The expectation is taken with respect to a subjective probability distribution. The general form is given, but not solved due to its intractability. A single-area example is used to demonstrate the technique. In order to obtain a solution, some constraints have to be applied to the probability distribution function.

Section 4 considers the problem of Knightian uncertainty: the existence of events whose probabilities cannot be ascertained beforehand or which are even unknown to the projector. The concepts of nonadditive utility and uncertainty aversion are introduced and used to motivate the solution of the problem. Their presence affects the solution. The single-area example of Section 3 is used as the basis of a numerical example.

Section 5 concludes this paper.

2. Loss Functions

Loss functions measure the "badness" of the departure of a projection from its actual value. The total loss function for a set of projections is

$$\mathcal{L} = \sum_{i=1}^{n} L(P_i; A_i) \equiv \sum_{i=1}^{n} \ell(\varepsilon_i, A_i) \qquad (1)$$

where i indexes the n areas projected, P_i and A_i are the projected and actual populations for area i, $\varepsilon_i = |P_i - A_i|$ is the absolute value of the projection error, and L and ℓ are the individual loss functions. In all cases, P_i and A_i are assumed positive. $\mathcal{L}$ is taken to be additive in order to satisfy the von Neumann-Morgenstern expected utility axioms. (Coleman, 2000a and 2000b) A total loss function which satisfies the von Neumann-Morgenstern axioms has the useful, if clumsily stated in this context, property that the loss associated with a gamble is equal to the probability-weighted sum of the losses.[1]

The individual loss functions are built by assuming an impartial decision-maker who has preferences over outcomes. The assumptions needed to create these functions are summarized below. For a fuller explanation, see Coleman (2000a). Subscripts are dropped, as they are not needed.

Assumption 1 (symmetry): $L(A + \varepsilon; A) = L(A - \varepsilon; A)$ for all $A > 0$.

Assumption 2 (monotonicity in error): $\partial\ell/\partial\varepsilon > 0$ for all $\varepsilon > 0$.

Assumption 3 (monotonicity in actual value): $\partial\ell/\partial A < 0$ for all $A > 0$.

Assumption 1 is very strong, as it implies that the decision-maker is indifferent between positive and negative errors. Assumption 2 simply states that smaller errors are preferred to larger ones. Assumption 3 states that an error of a given magnitude in a small area is worse than the same error in a large area. This can be best understood using an example. Suppose the error is 500. This is a serious error when the true value is 1,000, but almost a rounding error when the true value is 1,000,000.

The simplest loss functions that satisfy Assumptions 1-3 are:

$$L(P, A) = |P - A|^p A^q \qquad (2a)$$

[1]See von Neumann and Morgenstern (1944) for a statement of the axioms and the proof of this statement in terms of expected utility. Markowitz (1959, chap. 10) has an amended version of the von Neumann-Morgenstern axioms.

and

$$\ell(\varepsilon, A) = \varepsilon^p A^q \tag{2b}$$

where $\varepsilon, p > 0$ and $q < 0$.

Finally, several mathematical and statistical reasons exist to explain why absolute percentage errors decrease in the size of the area. To handle this, we assume Property 1:

Property 1: The loss function defined by equations (2a) and (2b) increases in A for any given absolute percentage error. This is assured whenever $q > -p$, or, equivalently, $p + q > 0$.

2.1 Example of Evaluating Population Projections Using Loss Functions

Loss functions can produce entirely different and more meaningful results than common error measures such as the mean absolute percentage error (MAPE). Table 1 at the end of this article shows the true values of six areas, A_i, $i = 1,\ldots,6$, and three sets (Scenarios) of absolute errors (ε_i), along with the corresponding absolute percentage errors (APE_i) and Webster's Rule loss function values (L_i). The bottom row shows the means of the last two variables. These are simply MAPE and $\mathcal{L}/n$, respectively. Webster's Rule sets $p = 2$ and $q = -1$. (Spencer, 1985) It is motivated by taking the view that projections are analogous to apportionments. (Coleman, 2000c) Balinski and Young (1982) found that Webster's Rule best satifies a large number of fairness criteria.

The three Scenarios are used to compare the results of an evaluation using a loss function to those obtained by using MAPE. Scenario 1 is the baseline scenario with $\text{APE}_i \equiv 2$ and $\mathcal{L}/n = 11.08$. In Scenario 2, APE_i is reduced to 1 for $i \le 5$, but APE_6 increases to 10. That is, all but the smallest areas have their APEs halved, but the very smallest area's APE increases by a factor of 5. MAPE increases to 2.5, but $\mathcal{L}/n$ falls to 2.9. Thus, MAPE ranks Scenario 2 as being less accurate than scenario 1, even though the individual errors are smaller except for the very smallest area. On the other hand, the loss function takes into account the size of the smallest area and discounts its accuracy loss and considers Scenario 2 to be more accurate. In Scenario 3, APE_i falls by 15% to 1.7 for $2 \le i \le 5$, rise by 50% to 3 in the largest area, and is unchanged in the smallest area. MAPE falls slightly to 1.97, but $\mathcal{L}/n$ rises to 18.19. Thus, MAPE considers Scenario 3 to be superior to Scenario 1, as a result of the general reduction in the APE_i, in spite of the major loss in accuracy in the largest area. The loss function, on the other hand, puts a large weight on the accuracy loss in area 1 and increases its error measure relative to Scenario 1. Thus, the loss function puts increasing weight on an error as the size of the area increases. Putting all of these together, we find that MAPE and the Webster's Rule loss function produce exactly opposite rankings of the Scenarios.

3. The Expected Total Loss Function

Assume that the joint subjective (Savage, 1954) distribution of the actual values is given by the Lebesgue-measurable probability density function $dF(A_1,\ldots,A_n)$. That is, the subjective probabilities associated with the actual values obey the customary laws of probability. We thus are dealing with "risk" in Knight's (1921) sense, in that the uncertainties are quantifiable. They are subjective in that they exist only in the mind of the projector. The future is unknowable, but the projector can make an estimate of dF. This estimate itself is based on a von Neumann-Morgenstern utility function on lotteries on all real n-tuples $(A_1,\ldots,A_n)$. (Anscombe and Aumann, 1963) The subjective expected total loss associated with a point forecast is the Lebesgue-Stieltjes integral

$$\mathbf{E}\mathcal{L} = \int_{\mathbf{A}} \sum_{i=1}^{n} L(P_i, A_i)\,dF \tag{3}$$

where $\mathbf{A}$ is the set of all real n-tuples $(A_1,\ldots,A_n)$.[2]

The objective of projection optimization is to choose a point projection $\mathbf{P} = (P_1,\ldots,P_n)$ to minimize $\mathbf{E}\mathcal{L}$, given dF.[3] This paper does a simple one area example to illustrate the problem.

Assume that a projection is made for one area at one point in time. Further, assume that Webster's Rule is used for the loss function. Then, the problem is to choose P^* to minimize

$$\begin{aligned} \mathbf{E}L(P,A) &= \int_{\underline{A}}^{\overline{A}} (P-A)^2 A^{-1}\,dF(A) \\ &= \int_{\underline{A}}^{\overline{A}} (A^{-1}P^2 - 2AP + A)\,dF(A) \end{aligned} \tag{4}$$

where $\underline{A}$ and $\overline{A}$ are the bounds of the support of $dF(A)$. To simplify matters, assume that $dF(A) = f(A)dA$ has a triangular distribution with mode A^*, $\underline{A} \le A^* \le \overline{A}$:

[2] For all infeasible A, $dF = 0$. These include all vectors with at least one impermissible projection value, such as a negative.

[3] Minimizing expected loss is equivalent to maximizing expected utility. (Coleman, 2000b) This does not lead to a circularity, as different utility functions are involved. The first utility function is applied to lotteries to obtain subjective probabilities. When the probabilities are objective, say, as the outcomes of spins of fair roulette wheels, the derived subjective probabilities are identical to the objective ones. (Anscombe and Aumann, 1963, p. 203) The second utility function is based on an individual's assessment of the outcomes (Coleman, 2000a) or on other normative criteria (Coleman, 2000c), which lead to Webster's Rule, used throughout the rest of this article.

$$
f(A)=\begin{cases} \dfrac{2(A-\underline{A})}{(A^*-\underline{A})(\overline{A}-\underline{A})} & \underline{A}\le A\le A^* \\[2ex] \dfrac{2(\overline{A}-A)}{(\overline{A}-A^*)(\overline{A}-\underline{A})} & A^*\le A\le \overline{A} \\[2ex] 0 & \text{otherwise.} \end{cases} \tag{5}
$$

The optimization problem now consists of substituting $f(A)dA$ from equation (5) for $dF(A)$ in equation (4) and finding the minimizing P^*.

The optimal P^* satisfies the equation

$$
P^*=\frac{(\overline{A}-\underline{A})(\overline{A}-A^*)(A^*-\underline{A})}{2\begin{bmatrix}\underline{A}A^*(\log A^*-\log\underline{A})\\ +\overline{A}A^*(\log\overline{A}-\log A^*)\\ +\overline{A}\underline{A}(\log\overline{A}-\log\underline{A})\end{bmatrix}}. \tag{6}
$$

Note that this is not a simple statistic, such as a mean or median. Its form exemplifies a general rule: P^* is, in general, a function of both the loss function and the underlying subjective probability distribution function.

4. Knightian Uncertainty

Section 3 assumed that the subjective probability density function dF was quantifiable. Since dF is subjective, there may remain residual uncertainty about its form. Moreover, dF does not take into account events whose probabilities are unknown. These events include those which cannot be foreseen altogether. Knight (1921) referred to this type of uncertainty as "uncertainty" itself. This now is frequently called "Knightian uncertainty." The upshot of the treatment and example used herein is that the presence of any Knightian uncertainty changes the loss-minimizing point projection.

Several methods exist for handling Knightian uncertainty, of varying usefulness for different applications. (Walley, 1999) The method used in this paper is Choquet capacities, which give rise to the Choquet integral. (Choquet, 1953) At the heart of this method is the concept of nonadditive probability. That is, given two events X and Y,

$$
\Pr(X)+\Pr(Y)\le \Pr(X\cup Y)+\Pr(X\cap Y). \tag{7}
$$

This is in contrast to the usual concept of Lebesgue-measurable probability, in which the inequality in (7) is replaced by an equality. It should be noted that the probability of the entire event space remains 1. For any given event X and probability density function dF, uncertainty aversion can be defined by

$$
c(dF,X)=1-\Pr(X)-\Pr(X^c) \tag{8}
$$

where X^c is the complement of X in the event space. "This number measures the amount of probability 'lost' by the presence of uncertainty aversion."[4] The "lost" probability reflects both the projector's ignorance over future events and his aversion to bearing uncertainty.[5]

The simplest assumption is constant uncertainty aversion.[6] Letting c be the uncertainty aversion, the corresponding Choquet capacity is $dF_c = (1 - c)\, dF$. Using the Choquet integral, Dow and Werlang (1992, p. 202) show that $\mathbf{E}_c\mathcal{L}$, the expected total loss which incorporates uncertainty aversion c, is given by[7]

$$
\mathbf{E}_c\mathcal{L}=c\sup_{\mathbf{A}}\mathcal{L}+(1-c)\mathbf{E}\mathcal{L}. \tag{9}
$$

The case $c = 0$ corresponds to complete certainty over dF and reduces $\mathbf{E}_c\mathcal{L}$ to $\mathbf{E}\mathcal{L}$. When $c = 1$, the projector has complete uncertainty aversion and sets his expected loss to be the maximum possible. In essence, his expected loss is his worst-case scenario. This scenario will be on the boundary of $\mathbf{A}$. He will choose a point estimate which minimizes his maximum total loss. That is, he will exhibit maximin behavior.[8] This point is further explored in Subsection 4.1. Intermediate values reflect the projector's possession of incomplete information about the future. In this case, $\mathbf{E}_c\mathcal{L}$ is a weighted combination of $\mathbf{E}\mathcal{L}$ and the worst-case loss. Thus, the loss-minimizing projection is intermediate between the two polar cases and is studied in Subsection 4.2.

4.1 Maximin Behavior

This is best exemplified by a one area problem. Using the notation of Section 3, when $c = 1$, the choice problem becomes to choose P^* to minimize

$$
\max_{[\underline{A},\overline{A}]}\left[L(P^*,\underline{A}),L(P^*,\overline{A})\right]. \tag{10}
$$

Given a loss function which obeys Assumption 1, P^* solves

$$
L(P^*,\underline{A})=L(P^*,\overline{A}). \tag{11}
$$

[4] Dow and Werlang (1992, p.200).

[5] See Schmeidler (1989, p. 582) for a formal definition of uncertainty aversion.

[6] Constant uncertainty aversion is a convenient assumption, but is not necessarily satisfied in reality.

[7] In terms of Dow and Werlang (1995), this is really $-\mathbf{E}_c(-\mathcal{L})$. The difference is that Dow and Werlang, (1995) Example 4.7, is concerned with maximization, while this problem is one of minimization.

[8] Strictly speaking, the projector exhibits minimax behavior with regard to expected total loss. However, per footnote 2, this is equivalent to maximizing minimum utility. Thus, it is appropriate to speak of the "maximin" rule. This rule was first proposed by Wald (1950) for decision-making in the presence of complete uncertainty. Ellsberg (1961) and Rawls (1971) subsequently proposed this rule for complete uncertainty in lotteries and the "initial position" of the wealth distribution of a society, respectively.

This equation results because both $\underline{A}$ and $\overline{A}$ are worst-case scenarios. Divergence from equality increases the loss with regard to one of $\underline{A}$, $\overline{A}$; thereby increasing the maximum loss. For Webster's Rule, equation (10) is solved by the geometric mean of $\underline{A}$ and $\overline{A}$. That is,

$$P^* = \sqrt{\underline{A}\overline{A}}. \tag{12}$$

This can be generalized to n areas.

4.2 Intermediate Uncertainty Aversion

This is best illustrated using the one area example of Section 3. Using Webster's Rule for L, the optimization problem is to choose P^* to minimize equation (9), given the EL of equation (4) and the probability density function $f(A)$ of equation (5). This problem requires a grid search over values of P until the expected loss-minimizing P^* is found. Figures 1a and 2a show the highly skewed $f(A)$ used to obtain the P^* and $E_c\mathcal{L}$ shown in Figures 1b and 2b, respectively. Problem (9) is solved for values of c ranging from 0 to 1 in increments of .1, as shown on the horizontal axes of Figures 1b and 2b. $c = 0$ is the no uncertainty aversion case solved in Section 3. When $c = 1$, solution (12) is obtained.

In both Figures 1a and 2a, $\underline{A} = 10$ and $\overline{A} = 30$. $A^* = 11$ and 29 in Figures 1a and 2a, respectively. Figures 1b and 2b show rapid convergence of P^* to $17.3 \approx \sqrt{300} = \sqrt{\underline{A}\overline{A}}$ as c increases. In both cases, $E_c\mathcal{L}$ increases in c. This is to be expected, as P^* converges to the maximin solution, which produces the greatest expected loss. The faster convergence to the maximin solution in Figure 1b appears to be because its P^* when $c = 0$ is closer to the maximin solution than that of Figure 2b.

Figures 1b and 2b both show that $E_c\mathcal{L}$ rises in c. The maximin solution represents the worst possible outcome, while the no uncertainty solution is able to use the subjective probability distribution to minimize expected loss. Intermediate values of c represent tradeoffs between the two. As c rises, P^* departs from the no uncertainty solution, thereby raising $E\mathcal{L}$. At the same time, greater weight is placed on the maximum loss. These effects together explain the rising $E_c\mathcal{L}$.

Figure 1a

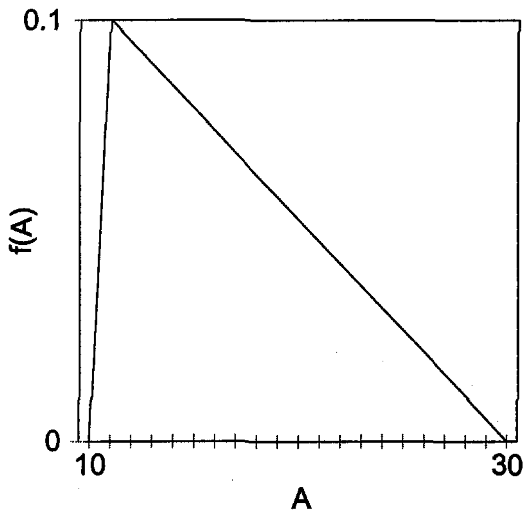


Figure 1b

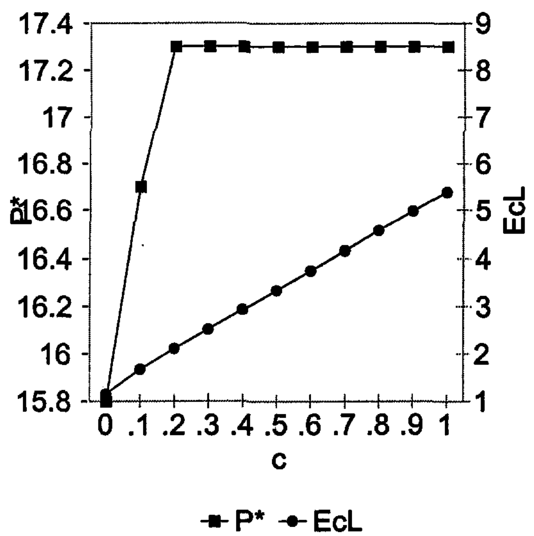


Figure 2a

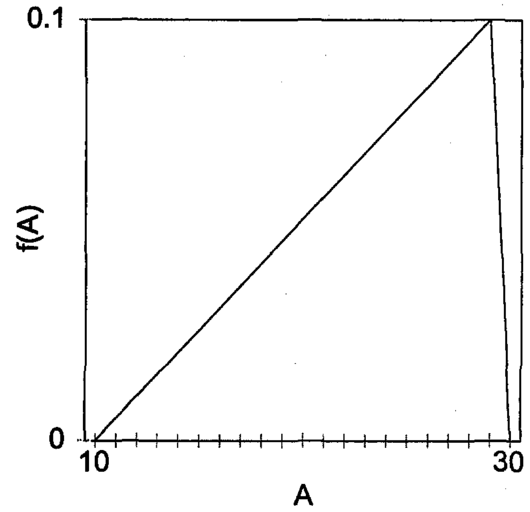
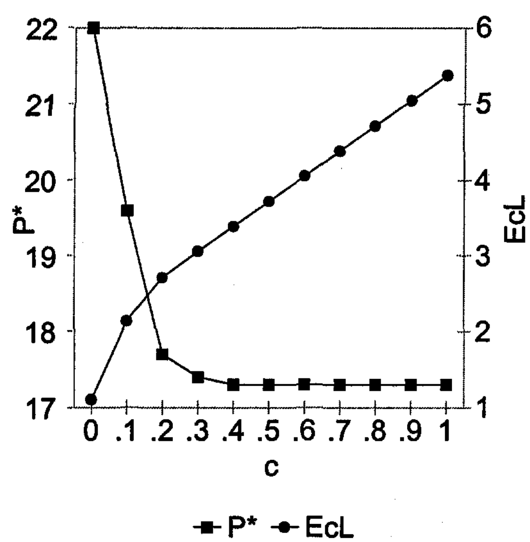


Figure 2b



5. Conclusion

This paper has considered the problem of creating point projections of population in order to minimize their expected total loss, given a subjective probability density function. The projection actually made is, in general, a function of both the loss function and the underlying subjective probabilities. Knightian uncertainty exists when there is residual uncertainty about the form of the subjective probability density function or when there exist events whose probabilities are unquantifiable or which may simply be unforseeable. A simple, constant uncertainty aversion model has been created for this case. If the projector has complete uncertainty aversion, he will select the maximin solution to minimize his maximum possible loss. Intermediate cases of uncertainty aversion result in projections intermediate between the zero uncertainty aversion case and the maximin solution. Expected loss rises in uncertainty aversion.

Although this paper is written in terms of cross-sectional population projections, its results are applicable to all manner of cross-sectional forecasts when the data are positive. The basic ideas remain the same. The methodology can be generalized to nonpositive data.[9]

6. Acknowledgements and Disclaimer

I would like to thank Ching-Li Wang for inspiring me to write this paper and thank him and Thomas Bryan for providing helpful comments.

This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a more limited review than official Census Bureau Publications. This report is released to inform interested parties of research and to encourage discussion.

References

Anscombe, F.J. and Aumann, R.J., (1963) "A Definition of Subjective Probability," *The Annals of Mathematical Statistics*, **34**, 199-205.

Bryan, Thomas, (1999) "U.S. Census Bureau Population Estimates and Evaluation with Loss Functions," *Statistics in Transition*, **4**.

Choquet, Gustave, (1953) "Theory of Capacities," *Annales de l'Institut Fourier*, **5**, 131-295.

Coleman, Charles D., (2000a) "Metrics for Assessing the Accuracy of Cross-Sectional Estimates and Forecasts," manuscript, U.S. Census Bureau. Previous version presented to 1999 Meetings of the Southern Demographic Association, San Antonio, TX, October 1999.

Coleman, Charles D., (2000b) "Total Loss Functions for Assessing the Accuracy of Cross-Sectional Estimates and Forecasts," manuscript, U.S. Census Bureau.

Coleman, Charles D., (2000c) "Webster's Rule for Assessing the Accuracy of Cross-Sectional Estimates and Forecasts," manuscript, U.S. Census Bureau.

Coleman, Charles D. and Bryan, Thomas, (2000) "Loss Functions for Detecting Outliers in Panel Data when the Data May Change Sign," manuscript, U.S. Census Bureau.

Dow, James and Werlang, Sérgio Ribeiro da Costa, (1992) "Uncertainty Aversion, Risk Aversion, and the

[9]Coleman and Bryan (2000) extend the loss function framework to nonpositive data, in the context of outlier detection.

Optimal Choice of Portfolio," *Econometrica*, **60**, 197-204.

Ellsberg, Daniel, (1961) "Risk, Ambiguity and the Savage Axioms," *Quarterly Journal of Economics*, **75**, 643-669.

Knight, Frank, (1921) *Risk, Uncertainty and Profit.* Boston: Houghton-Mifflin.

Markowitz, Harry M., (1959) *Portfolio Selection: Efficient Diversification of Investments*, New York: Wiley. Second edition, 1970: New Haven, CT: Yale Univesity Press. Third edition, 1991: Cambridge, MA: Blackwell.

von Neumann, John and Morgenstern, Oskar, (1944) *Theory of Games and Behavior*, Princeton University Press, Princeton. Second edition, 1947; third edition 1953: Princeton University Press, Princeton. Reprinted 1963 and 1967: John Wiley and Sons, New York.

Rawls, John, (1971) *A Theory of Justice*, Cambridge: Harvard University Press.

Schmeidler, David, (1989) "Subjective Probability and Expected Utility without Additivity," *Econometrica*, **57**, 571-587.

Spencer, Bruce D., (1985) "Statistical Aspects of Equitable Apportionment," *Journal of the American Statistical Association*, **80**, 815-822.

Wald, Abraham, (1950) *Statistical Decision Functions*, New York: John Wiley and Sons.

Walley, Peter, (1999) "Towards a Unified Theory of Imprecise Probability," *Proceedings of the 1st International Symposium on Imprecise Probabilities and Their Applications*, Ghent, Belgium, June 29 - July 2, http://decsai.ugr.es/~smc/isipta99/proc/055.html.

Table 1

		Scenario 1			Scenario 2			Scenario 3		
Area	A_i	ε_i	APE_i	L_i	ε_i	APE_i	L_i	ε_i	APE_i	L_i
1	100,000	2,000	2	40.00	1,000	1.0	10.0	3,000	3.00	90.00
2	50,000	1,000	2	20.00	500	1.0	5.0	850	1.70	14.45
3	10,000	200	2	4.00	100	1.0	1.0	170	1.70	2.89
4	5,000	100	2	2.00	50	1.0	0.5	85	1.70	1.45
5	1,000	20	2	0.40	10	1.0	0.1	17	1.70	0.29
6	100	2	2	0.04	10	10.0	1.0	2	2.00	0.04
Means			2	11.08		2.5	2.9		1.97	18.19

List of variables:

i: Area number
A_i: Actual value for area i
ε_i: Absolute error for area i
APE_i: Absolute percentage error for area i
L_i: Webster's Rule loss function value for area i

Projections of the Number of Households and Families in the United States: 1999 to 2025

by

Ching-li Wang

Population Projections Branch
Population Division
U.S. Census Bureau

I. Introduction

It has been a long history of preparing household projections in the U.S. Census Bureau. A new series of household projections is scheduled to release later this year. The purpose of this paper is to report Census Bureau's household projections and the methodology used in the current household projections from 1999 to 2025. Since future households depend on size of future population and its composition, this paper also discusses the effects of demographic trends and population projections on the results of household projections.

Household and family are the basic social and economic units of a society. A "household' is a person or group of people who occupy a housing unit. The householder is usually the person in whose name the housing unit is owned or rented. Households can be classified into two groups - family households and non-family households. A"family" is make up of two or more people living together who are related by blood, marriage, or adoption, and one of them is designated as the householder. Family households can be classified into several types based on marital status and presence of children - married couple family or other families with a female or male householder with no spouse present. In non-family households, people may live alone as a householder or with someone unrelated to the householder as in family households.

The projected number of families in this paper refers to the number of <u>family households.</u> Within a family household, there may be one or more families as subfamily without own household. Due to the limitation of methodology to project households based on marital status and householder rates, the current projections only project the number of family households by type. Therefore, the family used in the paper refers to family household.

Households and families provide the basic settings for living arrangements of population in a society. Public and private organizations use household and family statistics for policy and program development and implementation. The projections of the number and types of households and families into the future provide the information for such policy and program development.

II. History of household projections

The U.S. Census Bureau has a long history of preparing the household projections to meet these needs (See Appendix I). The earliest date the Census Bureau produced household projections can be traced back to 1943, when Paul Glick first estimated and projected households between 1940 and 1960. After revising the projections in 1946, he and other demographers in the Census Bureau did three projections in the 50s, and one in 1963 - projecting the number of households to year 1980.

Then Grymes and Parke took over the household projections in 1967. They also produced the household projections for states in 1968. That is the only time the Census Bureau released state household projections and was hoping to generate the interest of producing state household projections.

Since then, Grymes had been involved in the household projections activities for two decades. During this period of time, Jacob S. Seigel, Arthur J Norton and Donald J. Hernandez were also get involved. They updated three times in the 60s, three times in the 70s, and once in the 80s with last projections in 1986. After 1986, no one updated the projections until 10 years later. Jennifer Day prepared the last version of household projections in 1996. Four years later in the beginning of the new millennium, we are now finishing a new set of projections to be released later in 2000.

The procedures and steps to prepare the household projections have been changed slightly overtime since Glick initiated the program. In every new series, the marital status and householder rates were projected in different ways or new breakdown of household types and age groups were added. The most dramatic change of the methods to prepare household projections was in 1975 when Grymes began to use log transformation of householder rates to project the number of households and families.

III. Current household projections

The current projections project the number and type of households and families from 1999 to 2025. These projections are consistent with the 1990 census, as enumerated. They are not comparable with any post-1990 estimates of households from the Current Population Survey (CPS), which have been adjusted to include the net census undercount of approximately 4 million people.

The current household projections include the numbers and types of households and families, the average size of households and families, marital status of the population, and the number of families with children under 18 as prepared in previous projections (Day, 1996). In addition, the current projections include the number of households with adult children and living arrangements of people 65 and over.

As in previous series of projections, the current version of household projections include five basic *household types between 1999 and 2025.*

- o Family households
  1. Married couple family households
  2. Male householder family households with no spouse present
  3. Female householder family households with no spouse present

- o Non-family households
  4. Male householder
  5. Female householder

In these basic household types, additional subcategories are also given. -

- o Family households
  1. With children under 18
  2. *With no children under 18*
  3. With adult children 18 and over

- o Non-family households
  1. Living alone
  2. Living with other unrelated individuals

The age and sex of householders by types of households and families are also included in the projections.

The list of household and family types above represents two important variables of the population - marital status and householder rates. Therefore, the major task of household projections is focused on the assumptions and projections of marital status and householder rates in the future. The methodology used to create current projections is similar to the log-linear trend modeling of marital status and householder rates first used in 1975 by Gremyer (P25-607) and adopted in later versions of the household projections.

IV. Methodology

The number of households and families is a function of a population and its composition. The first step of projecting the number of households is to project the population or derive a set of population projections from the existing source. The second step is to project future marital and householder rates by type of households. Then the projected marital and householder rates are applied to the population projections to derive the projected number of households and families. The middle series of most recent U.S. population projections to year 2100 were used (U.S. Census Bureau, Working Paper No. 38, January, 2000).

(a). Marital status proportions and householder rates

Marital status and householder rates were computed from the 1990 census and 1959-1998 CPS data. The proportions of never married and ever married household population were calculated by sex (male and female) and up to 12 age groups (15-17, 18-19, 20-24, 25-29, 30-34, 35-44, 45-54, 55-59, 60-64, 65-74, 75-84, and 85+) for the CPS time series data. The marital status and householder rates by race/origin (non-Hispanic White, non-Hispanic Black, non-Hispanic American

Indian or Alaska Native, non-Hispanic Asian or Pacific Islander, and Hispanic of any race) from the 1990 census are also used. The proportions of ever married persons who were married, spouse present, were calculated for the same age, sex, race, and Hispanic origin groups. A list of marital and householder rates used in the projections is shown in Appendix I.

(b). Projection of marital status and householder rates

The projections of marital status and householder rates were based on time series data from the CPS (Current Population Surveys). The marital status (never married, married with spouse present, and married males with own households) by age and sex were projected 27 years to the year 2025 based on the data between 1959 and 1998. The householder rates to derive household types (family households, families with no spouse present, and non-family households) by age and sex of householders were based on the CPS data from 1969 to 1998. The choice of 1959 or 1969 as the starting points of the time series was based on which showed the smoothest and most consistent trends. Outliers and proportions of zero or one were omitted from the time series for projections of rates. The CPS data required several modifications in order to preserve a consistent series of householder rates (Current Population Reports, P25-986).

(1). Logistic transformation of marital status and householder rates

As in earlier household projection series (P25-805), the average annual changes in the CPS time series data were used to create the future marital status and householder proportions for the year 2025. The marital status and householder rates from the CPS data were log transformed in order to better approximate normal distributions. Where x_t is a proportion in year t, the transformed value, y_t, is:

$$y_t = \log(x_t / (1 - x_t))$$

In addition, the natural log of proportions, $\log(x_t)$ was taken before the regression to prevent the projected proportions from going below zero. The natural log of proportions, logarithm one minus the proportion, $\log(1-x_t)$ was taken to prevent projected proportions from going above one. The two factors are combined, $\log(x_t/(1-x_t))$, to prevent projected proportions from going below 0 or above 1 (Bell, Bozik, McKenzie, and Shulman, 1986).

(2). Projection of transformed rates to year 2025

The linear least squares of the log transformed 1959-1998 or 1969-1998 marital status and household rates are used to estimate the annual change of rates, which were applied to the 1998 starting points, by age, sex, race, and origin, and were projected 27 years into the future as the following formula.

$$y_{(t:2025)} = y_{(t:1998)} + \text{LNEST}(y{:}1959\text{-}1998)*27$$

(3). Inverse transformation of projected value

The projected values of y_t were used to forecast the values of x_t in 2025 with the inverse transformation as follows:

$$x_t = \exp(y_t)/(1+\exp(y_t))$$

(4). Linear interpolation of projected rates between 1998 and 2025

The householder rates for the years between 1998 and 2025 were linearly interpolated to generate a smooth line.

$$x_t = x_{t(1998)} + (x_{t(2025)} - x_{t(1998)})* (n/27)$$

Where, n is the number of years from 1998.

(C). Three series of projections of marital status and householder rates

Three series of projections were prepared based on the different assumptions of future marital status and householder rates.

Series 1, Adjusted Trend Projections:

Series 1 is initially based on a log-linear extrapolation of changes of marital status and householder rates derived from the Current Population Survey (CPS) from 1959 to 1998 or 1969 to 1998. Then these extrapolated rates were adjusted to reflect the assumption that the demographic changes affecting household and family formation would slow during the projection period.

Various demographic factors influence the number and types of households. Age at first marriage influences the proportion of people never married. Increased age at first marriage can lead to an increase in the proportion of younger persons in non-family living arrangements, either living alone or with roommates, and can reduce the proportion of persons maintaining family households. Divorces can influence household composition by leading to increases in adults forming their own households, family households with no spouse and non-family households, thereby reducing the proportion maintaining married couple households. Nonmarital childbearing increases the proportion of family households with children.

Many of the demographic factors described above changed dramatically during the 1970s and 1980s. More recently, some of these demographic changes have slowed and, in some cases, reversed themselves[1]. Therefore, it is assumed that age at first marriage will continue to rise, but at a slower pace in the future. The divorce rate declined slightly after 1979. The leveling of divorce also moderates change in the proportion of people with children but no spouse in the home, especially for women. The proportion of men maintaining families without spouses present has been increasing and will continue to increase.

As with the previous household projections (Current Population Reports, P25-1129), some adjustments of projected marital and householder rates were made before the projections of households and families. The projected changes of proportions of persons who are never married between 1998 and 2025 were reduced by 3/4 for males and females of all ages. The 27 year decline in the proportion of males and females who are married, spouse present was reduced by 2/3 for all ages. Finally, the projected changes between 1998 and 2025 in the proportions of male and female family householders with no spouses present and male and female non-family householders were reduced by 2/3 for all ages.

Series 2: Historical Trends Projections

Series 2 is also based on a log-linear extrapolation of marital status and householder rate changes in the CPS as Series 1, but with no adjustments of current trends. This series simply accepts the trends based on original time series data from 1959 or 1969 without any adjustment for current trends. So, Series 2 illustrates the impact of continuation of marital and household trends in the past on the number and types of future households and families.

Series 3: Constant Rates Projections

For comparative purposes, the marital status and householder proportions by age, sex, race, and Hispanic origin from the 1990 Census were held constant to project the number of households and families as Series 3. This series shows only the effects of the projected changes in the demographic structure of the population without any changes in household and family formation rates.

(D). Procedures to derive projected households and families

(1). Preparation of starting point estimates

The 1998 household estimates by type are used as the starting point for the household projections. Because detailed household estimates by type, consistent with the 1990 Census are not available from the current household estimates prepared by the Census Bureau, the creation of the 1998 household estimates required several steps. First, the household populations for each year 1990 to 1998 was prepared by applying the 1990 census proportions of the resident population living in group quarters by age, sex, and race/ethnicity to resident population.

The 1998 detailed household estimates were calculated by applying the annual proportional changes in marital status and householder rates from the 1990-98 CPS to the corresponding 1990 census data. The marital status proportions were applied to the estimated household population to derive the estimates of the number of people who were ever married and married with spouse present. Householder rates were applied to the appropriate marital status groups to generate the estimates of households by type and age, sex, race, and origin of the householders for 1990 through 1998. These

[1] For further information, see P20-514 and Lynne M Casper and Ken Bryson. 1998. Household and Family Characteristics: March 1998 (Update), P20-515. U.S. Government Printing Office, Washington, DC.

estimates were controlled to agree with the official household estimates prepared by the Census Bureau (ST-98-46).

The difference between the total numbers of households, controlled to the previous estimates (1990 through 1998) and the projected number of households with no controls produced shifts in the rate of household change from 1998 to 1999. Therefore, an average of the 1990-1998 control factors, by age group, was applied to the total projected number of households for every projection year from 1999 to 2025.

(2). Creation of the projected household population

The middle series of the most current resident population projections for the U.S. was used to derive a household population (Census Bureau, Working Paper No. 38, January 2000). The projections include people living in institutions, non-institutional group quarters (such as college dormitories and military quarters), and households. The 1999 proportions of non-institutional population by age, sex, race, and Hispanic origin were applied to each year of the middle series of population projections, 1999 to 2025. The 1990 proportions of people living in other group quarters were also applied to the population projections. The projected household population was computed by subtracting the projected number of people living in group quarters from the non-institutional population.

(3). Application of projected marital and householder rates

Follow the similar steps used in previous version of household projections (*Current Population Reports*, P25-805, May 1979), the number of never married males and females was first calculated by multiplying the projected proportions of males and females who were never married by the corresponding projected household population. The difference between the household population and the never married population is the projected ever married population.

From the projected ever married population, the currently married males and females with spouse present are calculated by applying the projected proportions of married males and females with spouse present. Since there must be an equal number of married, spouse present men and women, the preliminary total numbers of married, spouse present males and females were averaged for each projection year. A ratio of the average number of married, spouse present persons to the preliminary total number of married, spouse present males was calculated. This ratio was multiplied by the projected number of married, spouse present males by age, race, and origin to generate a proportionally adjusted number equal to the average number of married, spouse present persons. The same procedure was performed for females.

The projected number of married couple households was computed by multiplying the householder rate for married males with their own household by the adjusted number of married, spouse present males. Married couple households are represented by the husband's age, race, and Hispanic origin in order to simplify the calculations and tables.

The difference between the projected household population and the adjusted married couples population is the projected number of not currently married population which includes those who are not married or no spouse present. From the projected number of people who are not currently married, the number of family households with no spouse present is derived by applying the projected proportions of male and female family householders with no spouse present.

Projected non-family households are also derived from the number of not currently married males and females or married with no spouse present. The projected proportions of non-family male and female householders (or primary individuals) were multiplied by the number of not currently married males and females to project the number of non-family households.

The projected numbers of married couple families, other families (male and female family householders with no spouse present), and non-family households (male and female non-family householders) are adjusted by multiplying the average of the 1990-1998 control factors to produce total numbers of projected
households consistent with the official estimates.

Table 1. **Number of Households and Average Annual Increase: 1940 to 2025**

[In thousands. Reference date is July 1, except as noted]

Year	Number of households			annual change from previous date		
	Series 1	Series 2	Series 3	Series 1	Series 2	Series 3
CENSUS ESTIMATES						
1940*		34,949			(X)	
1950*		43,468			2.2	
1960*		52,610			1.9	
1970*		63,450			1.9	
1980*		80,390			2.4	
1990*		91,947			1.3	
1998**		101,041			1.2	
PRELIMINARY PROJECTIONS						
1999	102,426	102,681	101,822	1.4	1.6	0.8
2000	103,652	104,173	102,921	1.2	1.4	1.1
2005	109,783	111,758	108,401	1.1	1.4	1.0
2010	116,096	119,692	114,041	1.1	1.4	1.0
2015	122,412	127,737	119,682	1.0	1.3	0.9
2020	128,553	135,759	125,063	0.9	1.2	0.8
2025	134,647	144,063	130,206	0.9	1.2	0.8

*As of April 1, from population censuses.

** 1998 Census-based estimate

X Not Applicable

Sources: U.S. Bureau of the Census. *Historical Statistics of the United States, Colonial Times to 1970*, Bicentennial Edition, Part 2. Washington, DC, 1975, p. 42.; Census of the Population: 1970, Volume 1. *Characteristics of the Population, part 1, United States Summary*-Section 1. U.S. Government Printing Office, Washington, DC, 1973, p. 1-278.; 1980 Census of Population, PC80-1-B1, *United States Summary*. U.S. Government Printing Office, Washington, DC, 1983, p. 1-44.; 1990 Census of Population, General Population Characteristics, United States, 1990 CP-1-1, U.S. Government Printing Office, Washington, DC, 1992; and table 1.

V. Results and Discussion

According to Series 1, the number of households in the United States is projected to increase by over 32 million from 102.4 million in 1999 to 134.6 million in 2025 (Table 1).. This represents 31.5 percent increase or an average annual increase of 1.0 percent between 1999 and 2025, considerably slower than any historical period since 1940 This growth rate translates to an expected annual increase in the number of households between 1.2 and 1.3 million per year for the projection period – slightly higher than the average growth in number of households during the 1990s (Table 1).

Under Series 2, with the assumption of continuation of historical trend in the past four decades, the annual change between 1999 and 2025 in number of households ranges from 1.2 percent to 1.4 percent per year (Table 1) and results in 144 million households by 2025. This represents an increase of over 41 million or 40.3 percent from 1999 to 2025 and is 9 million more than projected in Series 1.

If householder rates were to remain constant at the 1990 census levels, as shown in Series 3, the increases in the expected number of households would be 28 million, smaller than projected in Series 1 and 2. Series 3 projects annual increases in the number of households of 1.02 to 1.16 million per year with average annual increases of 0.8 to 1.1 percent.

Under Series 3, the marital status and householder rates were held constant throughout the projection period. The increase in the number of households between *1999* and *2025 in* Series 3 can be attributed to the changes only in population and its age, sex, race and origin composition. The difference between Series 3 and Series 1 or 2 can be used to measure the effect of different assumptions of marital and household rates or household formation on the projected number of households. As table 2 shows, changes in age, sex, race, and origin account for a larger percent of changes in the projected number of households between 1999 and 2025. Changes in household formation account for only 12 percent of the projected increase in number of households under Series 1, and 31.4 percent under Series 2 between 1999 and 2025 (Table 2) This different is expected because the

Table 2. **Projected Increase in the Number of Households by Compositional Change:1999-2025**
[In thousands]

Year or period	Series 1	Series 2	Series 3
HOUSEHOLDS			
1999	102,426	102,681	101,822
2005	109,783	111,758	108,401
2010	116,096	119,692	114,041
2015	122,412	127,737	119,682
2020	128,553	135,759	125,063
2025	134,647	144,063	130,206
INCREASE BY COMPONENTS			
1999 to 2005, total	7,357	9,077	6,579
Age-sex-race-origin compositional change	6,579	6,579	6,579
Household formation changes	778	2,498	0
% due to household formation changes	10.6	27.5	-
2005 to 2010, total	6,312	7,935	5,640
Age-sex-race-origin compositional change	5,640	5,640	5,640
Household formation changes	672	2,295	0
	10.7	28.9	-
2010 to 2015, total	6,316	8,045	5,641
Age-sex-race-origin compositional change	5,641	5,641	5,641
Household formation changes	675	2,404	0
% due to household formation changes	10.7	29.9	-
2015 to 2020, total	6,142	8,021	5,381
Age-sex-race-origin compositional change	5,381	5,381	5,381
Household formation changes	761	2,641	0
% due to household formation changes	12.4	32.9	-
2020 to 2025, total	6,094	8,304	5,143
Age-sex-race-origin compositional change	5,143	5,143	5,143
Household formation changes	951	3,161	0
% due to household formation changes	15.6	38.1	-
1999 to 2025, total	32,221	41,382	28,384
Age-sex-race-origin compositional change	28,384	28,384	28,384
Household formation changes	3,837	12,998	0
% due to household formation changes	11.9	31.4	-

- Represents zero.

householder rates were adjusted under Series 1 and not adjusted under Series 2.

However, the percent of changes in number of households accounted for by changes in household formation rates increases from 10.7 percent in 1999-2005 to 15.6 percent in 2020-2025 under Series 1 projections. The percentage increases from 27.5 percent in 1999-2005 to 38 percent in 2020-2025 under Series 2 projections. In other words, although changes in household formation rates account for only a smaller percentage of changes in number of households, the percentage increases overtime.

Nevertheless, future population change is a key component of the projected number of households- accounts for 88 percent of changes in number of households between 1999 and 2025. The assumptions used to create the population projections (i.e. future fertility, mortality, and net immigration) determine much of the expected growth of the household population.

The age composition of the population is also an important component of household growth since most new households are established by young adults. As people move along their life course and transition to different types of households (such as through marriage, childbearing, divorce, or widowhood), the size of the cohort passing through each stage of life will affect the number of households and the type of households created in the process. Under Series 1, the projected slower growth in the total number of households is due to the relatively small cohorts of young adults who will be forming new households during the next 26 years. The large Baby Boom cohorts which affected dramatic growth of households in the 1970s and 1980s are moving toward middle age and completing their family formation. By 2025, all the baby boomers will be over 60. The majority of them will be in the empty nest stage – which will affect the household composition more than the total number of households.

Since the householder rates are applied to population projections, all of the assumptions about fertility, mortality, and migration incorporated into the population projections also affect the household projections. To illustrate the impact of population projections on household projections, the low population projection and high projections series produced by the U.S. Census Bureau (Working Paper No. 38) are used to project the number of households as shown in Table 3.

The application of Series 1 martial status and householder rates to the low alternative population projections series produces 127 million households by 2025, 7.8 million households less than projected using middle series of population projections (Table 3). The application of Series 1 marital status and householder rates to the high alternative population projections produces 146 million households in 2025, 11.8 million more than projected middle series. The difference between high and low projections based on the alternative population projections is much larger than the difference between Series 2 (unadjusted time series model) and Series 3 (constant rates model) - 20 million vs. 14 million. In other words, the use of alternative population projections produces larger difference in total projected number of households in the future.

However, when we examine the number of households by types produced by using the alternative population projections, and based on historical trends (Series 2) or constant rates (Series 3), the patterns are shift. The use of different householder rates by Series 2 and Series 3 produces wider range of difference than the use of alternative population projections in in all types of households. For example, Series 2 projects 55 million married couple family households in 2025, less than 15 million than projected by Series 3. The use of lower alternative of population projections projects 59 million married couple family households in 2025, 9 million less than the use of high alternative population projections. All other types of households were

Table 3. **Alternative Household Projections by Household Types Using Different Population Projections:1999, and 2025**
[In thousands]

Year and type	Series 1 Lowest	Series 1 Middle	Series 1 Highest	Series 2	Series 3	Difference Between High and Low Alterntives	Difference Between Series 2 and Series 3
Population Projectcions							
1999	272,695	272,820	272,957	272,820	272,820	262	-
2025	308,229	337,815	380,397	337,815	337,815	72,168	-
1999							
All households	**102,390**	**102,426**	**102,466**	**102,681**	**101,822**	76	858
Married couple family	52,712	52,730	52,750	52,408	56,287	37	-3,879
Female householder family	12,940	12,946	12,953	13,087	11,871	14	1,216
Male householder family	4,491	4,493	4,495	4,572	3,778	5	794
Female nonfamily	17,137	17,142	17,148	17,308	16,203	11	1,106
Male nonfamily	15,111	15,115	15,120	15,306	13,683	9	1,623
2025							
All households	**126,788**	**134,647**	**146,408**	**144,063**	**130,206**	19,620	13,857
Married couple family	58,867	62,381	67,656	54,729	69,708	8,789	-14,979
Female householder family	16,580	17,838	19,758	21,631	14,796	3,178	6,835
Male householder family	6,689	7,265	8,188	9,857	5,426	1,499	4,431
Female nonfamily	23,427	24,688	26,387	29,849	21,744	2,960	8,106
Male nonfamily	21,226	22,475	24,419	27,996	18,532	3,194	9,464
Percent Change 1999-2025							
All households	23.8	31.5	42.9	40.3	27.9	19.1	12.4
Married couple family	11.7	18.3	28.3	4.4	23.8	16.6	-19.4
Female householder family	28.1	37.8	52.5	65.3	24.6	24.4	40.6
Male householder family	48.9	61.7	82.1	115.6	43.6	33.2	72.0
Female nonfamily	36.7	44.0	53.9	72.5	34.2	17.2	38.3
Male nonfamily	40.5	48.7	61.5	82.9	35.4	21.0	47.5

Sourse:
U.S. Census Bureau Internet Release (January 13, 2000)
http://www.census.gov/population/projections/nation/summary/np-t1.txt
(NP-T1) Annual Projections of the Total Resident Population as of July 1:
Middle, Lowest, Highest, and Zero International Migration Series,
1999 to 2100.

projected with wider range of growth rates between Series 2 and Series 3 than the use alternative population projections as shown in Table 3.

Thus, it is evidence that the demographic changes determine the number of future households, but different assumptions of the changes in householder rates have very significant effects on the composition of households and families. Therefore, it is critical to have an appropriate population projections as the base for household projections and also important to have reasonable marital status and householder rates for projecting types of households and families.

Appendix I: Household projections prepared by U.S. Census Bureau

Day, Jennifer Cheeseman, "Projections of the Number of Households and Families in the United States: 1995 to 2010." U.S. Census Bureau, *Current Population Reports*, P25-1129, April 1996.

Grymes, Robert O. and Donald J. Hernandez, "Projections of the Number of Households and Families: 1986 to 2000." U.S. Census Bureau, *Current Population Reports*, P25-986, May 1986.

Grymes Robert O. and Arthur J. Norton, "Projections *of the Number of Households and Families:* 1979 to 1995." *Current Population Reports*, P25-805, May 1979

Grymes, Robert, "Projections of the Number of Households and Families: 1975 to 1990." U.S. Census Bureau, *Current Population Reports*, P25-607, August 1975.

Parke, Robert Jr. and Robert O. Grymes, "Projections of the Number of Households and Families: 1967 to 1985." U.S. Census Bureau, *Current Population Reports,* P25-394, June 1968.

Heer, David M. and Paul C. Glick, "Illustrative Projections of the Number of Households and Families: 1960 to 1980." U.S. Census Bureau, *Current Population Reports*, P20-90, October 1958.

Landau, Emanual, "Projections of the Number of Households and Families: 1960 to 1975." U.S. Census Bureau, *Current Population Reports*, P20-69, August, 1956.

Glick, Paul C., Elizabeth A. Larmon, and Emanual Landau. "Projections of the Number of Households and Families: 1955 to 1960," U.S. Census Bureau, *Current Population Reports,* P20-42, December 1952.

Glick, Paul C., "Estimated Number of Families in the United States: 1940 to 1960." U.S. Census Bureau, *Population - Special Reports*, Series P-46, No. 4, June, 1946.

Glick, Paul C., "Estimated Number of Families in the United States: 1940 to 1960." U.S. Census Bureau, *Population - Special Reports*, Series P-1943, No.2, September 1943.

Appendix II: Marital status and Householder rates used in the household projections

In past projections the denominator was sometimes the resident population. The "H" is to indicate that the denominator is the household population. All rates or proportions below were calculated by age groups and by race/origin.

1. Proportion of single female
   SFH = (number of never married females) / (number of females in the household pop)
   Note: the projected changes were reduced by *3/4 for series 1.*
2. Proportion of single male
   SMH = (number of never married males) / (number of males in the household pop)
   Note: the projected changes were reduced by 3/4 for series 1.
3. Proportion of married females with spouses present
   MFSPH = (number of married females with spouses present) / (number of ever married females)
   Note: the projected changes were reduced by 2/3 for series 1.
4. Proportion of married males with spouses present
   MMSPH = (number of married males with spouses present) / (number of ever married males)
   Note: the projected changes were reduced by 2/3 for series 1.
5. Proportion of married males with own households
   MMOHH = (number of married males with their own households) / (number of married, spouse present males)
   Note: the numerator omits married couples who are living in someone else's household; this is also where married couple households are represented by the male spouse's characteristics only.
6. Proportion of female family householders with no spouse
   FFH = (number of female family householders) / (number of not currently married females)
   Note: this is the rate for female householders *with one or more relatives in the household* and no spouse; projected changes were reduced by 2/3 for series 1.

7. Proportion of male family householders with no spouse
   MFH = (number of male family householders) / (number of not currently married males)
   Note: this is the rate for male householders with one or more relatives in the household and no spouse; projected changes were reduced by 2/3 for series 1.
8. Proportion of female non-family householders
   FPI = (number of female non-family householders) / (number of not currently married females)
   Note: this is the rate for females living alone or only with non-relatives; projected changes were reduced by 2/3 for series 1.
9. Proportion of male non-family householders
   MPI = (number of male non-family householders) / (number of not currently married males)
   Note: this is the rate for males living alone or only with non-relatives; projected change were reduced by 2/3 for series 1.

References

Bell, William, James Bozik, Sandra McKenzie, and Holly Shulman. 1986. "Time Series Analysis of Household Headship Proportions: 1959-1985." U.S. Census Bureau, Statistical Research Division Report, CENSUS/SRD/RR-86/01.

Day, Jennifer Cheeseman. 1996. Projections of the Number of Household and Families in the United States: 1995 to 2010. U.S. Census Bureau, Current Population Reports, P25-1129. Washington, D.C.: U.S. Government Printing Office.

Hernandez, Donald J. 1995a. "Review of Preliminary Household Projections." U.S. Census Bureau internal memo dated June 16, 1995.

Hernandez, Donald J. 1995b. "Review of July 12 Revised Household Projections." U.S. Census Bureau internal memo dated July 14, 1995.

Hollmann, Frederick, Tammany Mulder, and Jeffrey Kallan. 2000. Population Projections of the United States, 1999 to 2100: Methodology and Assumptions.U.S. Census Bureau, Population Division, Working Paper No. 38.

Siegel, Jacob S. 1972. "Development and Accuracy of Projections of Population and Households in the United States." Demography. 9(1): 51-68.

U.S. Census Bureau. 1975. Projections of the Number of Households and Families : 1975-1990. Current Population Reports, P25-706. Washington, D.C.: U.S. Government Printing Office

U.S. Census Bureau. 1977. Population Projections of the United States by Age, Sex, Race, and Hispanic Origin. Current Population Reports, P25-704. Washington, D.C.: U.S. Government Printing Office.

U.S. Census Bureau. 1979. Projections of the Number of Households and Families: 1979 to 1995. Current Population Reports, P25-805. Washington, D.C.: U.S. Government Printing Office.

U.S. Census Bureau. 1986. Projections of the Number of Households and Families: 1986 to 2000. Current Population Reports, P25-986. Washington, D.C.: U.S. Government Printing Office.

U.S. Census Bureau. 1996. Estimates of Housing Units, Households, Households by Age of Householder, and Persons per Household of States: Annual Time Series, July 1, 1990 to July 1, 1996, ST-96-24R, U.S. Government Printing Office, Washington, DC.

U.S. Census Bureau. 1999. U.S. Population Estimates by Age, Sex, Race, and Hispanic Origin: 1980 to 1998 (with extension to August 1, 1999). http://www.census.gov/population/www/estimates/nat_90s_detail.html, released October 1, 1999.

U.S. Census Bureau. 1999. Estimates of Housing Units, Households, Households by Age of Householder, and Persons per Household of States: Annual Time Series, July 1, 1990 to July 1, 1996, ST-98-46, U.S. Census Bureau Internet Release:http://www.census.gov/population/estimates/housing/sthuhh1.txt

WILL STRONG U.S. GROWTH CONTINUE? A LOOK AT U.S. GROWTH IN THE 1990s AND ITS IMPLICATIONS FOR THE U.S. GROWTH OUTLOOK

Chair: Paul Sundell
Economic Research Service, U.S. Department of Agriculture

Will Strong U.S. Growth Continue? A Look at U.S. Growth in the 1990's and Its Implications for the U.S. Growth Outlook—Abstract

Panelists:

Paul Sundell, Economic Research Service, U.S. Department of Agriculture

Robert W. Arnold, Congressional Budget Office

Ralph Monaco, INFORUM, University of Maryland

The U.S. Economic Outlook for 2001: Slower Growth Finally Arrives,
Paul Sundell, Economic Research Service, U.S. Department of Agriculture

The Outlook for Productivity Growth: Are We in a New Economy?,
Robert W. Arnold, Congressional Budget Office

Will Strong U.S. Growth Continue? A Look At U.S. Growth in the 1990's and Its Implications for the U.S. Growth Outlook

Chair: Paul Sundell, Economic Research Service, U.S. Department of Agriculture

The United States has experienced falling inflation and very rapid real economic growth since 1995. This situation has forced many economists to significantly raise their estimates of potential gross domestic product (GDP) and their forecasts for future U.S. productivity and real economic growth. The increase in estimated potential GDP and expected stronger future economic growth have been primarily attributed to three positive factors. First, a lower estimated NAIRU (nonaccelerating inflation rate of unemployment) implies an increased supply of labor available consistent with nonaccelerating inflation. Second, the business investment boom has increased the quantity and quality of capital available per worker. Increased capital per worker leads to an increase in potential GDP by both moving along the production function and shifting the production function outward over time by raising total factor productivity. Third, total factor productivity (TFP) has risen due to various favorable nonbusiness investment supply and demand factors. Positive non-investment supply factors include the increased globalization of the U.S. economy, more efficient management structure, and the falling real prices for energy, food, employee benefits, and imported goods. Total factor productivity has also been raised, at least temporarily, by strong growth in aggregate demand raising the intensity existing labor is utilized in the production process.

The questions now are: Will U.S. growth slow moderately or will U.S. growth slow dramatically to well under three percent for most of the future?; and How should economic performance and productivity best be measured? In this session these questions will be addressed by the panelists. Paul Sundell will discuss his research on measuring potential GDP and TFP growth, and the implications of his research for future U.S. productivity and economic growth. Robert W. Arnold will discuss his views concerning the U.S. growth and productivity outlook, emphasizing the near-term and intermediate-term (less than 5 year) outlook. Ralph Monaco will discuss his views concerning U.S. economic growth and productivity outlook emphasizing the intermediate-term to longer-term outlook (over 5 years). He will also examine problems in measuring productivity and review some recent attempts to characterize the economy's performance in ways other than traditional labor productivity or total factor productivity calculations.

Panelists:

Paul Sundell
Economic Research Service, U.S. Department of Agriculture

Robert W. Arnold
Congressional Budget Office

Ralph Monaco
INFORUM, University of Maryland

The U.S. Economic Outlook For 2001: Slower Growth Finally Arrives

Paul A. Sundell, Economic Research Service of USDA

Introduction

While real economic growth slowed to approximately 3.1 percent in the second half of 2000, year over year economic growth is still expected to average 5.1 percent for 2000 and 3.0 percent for 2001. The moderation in U.S. growth reflects numerous factors: (1) tight labor markets, (2) U.S. GDP exceeding non inflationary potential GDP, (3) much tighter conditions in business capital markets, (4) slower growth in consumer and residential housing spending, and (5) higher oil prices. The slowdown in economic growth to 3.0 percent in 2001 represents a return to more normal sustainable growth. In comparison, economic growth over the 1980-1999 period averaged 3.0 percent. U.S. economic growth will be sustained by: (1) strong underlying productivity growth, (2) above average (but slowing) growth in business fixed investment, and (3) strong foreign growth. Inflation, as measured by the GDP deflator, is expected to rise slightly to 2.4 percent due to the contemporaneous and lagged effects of tight labor markets that are slowly accelerating labor costs, lower productivity growth, and, to a lesser extent, higher energy and import prices.

The paper discusses each of these factors and its implications for the U.S. economic outlook. Special emphasis is placed on two areas: (1) measuring potential GDP and the impact of the existing gap between actual GDP and potential GDP on short-term economic growth and (2) the tightening of conditions in U.S. business capital markets and its impact on near term growth outlook. Potential GDP was estimated using a production function approach with total growth in total factor productivity varying both deterministicly across business cycles and stochastically over time. A time varying NAIRU series was estimated with the estimated NAIRU at 5.5 percent for 1999IV. The estimated NAIRU series was used as an input in deriving nonaccelerating inflation potential labor hours series used in constructing the potential GDP. GDP was estimated to have exceeded its potential by 3.7 percent by the end of 1999.

In the late 1990's and the first half of 2000, the normal slowing of economic growth that occurs when GDP exceeds its long-term potential was offset by large productivity gains, a boom in business fixed investment spending, and a near doubling of equity prices over the 1996-1999 period. The combination of falling import prices, food and energy prices, and slower growth in medical costs temporarily further boosted aggregate supply and enabled inflation to fall over most of the 1996-1999 period (Brinner, Rich and Rissmiller, and Browne pp.5-8). The relative price of imports, energy, and medical costs have moved upward in 2000 and will raise inflation slightly in 2001.

This paper discusses the significant tightening of conditions in business debt and equity markets in 2000. The combination of higher interest rates, tighter credit standards, and overall depressed equity markets will significantly slow the growth in business fixed investment in 2001. The tighter capital markets conditions will impact firms with poorer default and liquidity conditions relative to firms in stronger financial condition. Rising default premiums on corporate bonds and tightening lending standards on business loans have empirically been associated with slower economic growth (Duca, Van Horne, and Lown, Morgan and Rohatgi)

Major Factors Supporting Near Term Strong U.S. Economic Growth

Continuation of Strong Productivity Growth

Since 1995, labor productivity growth accelerated sharply to an annual rate of 2.9 percent compared to 1.6 percent annually over the 1991 to 1995 period. Normally, productivity growth slows down as an economic recovery matures. In this economic recovery, productivity growth has accelerated indicating that longer term trend productivity growth has likely moved upward as well. Productivity has been boosted since 1995 by numerous factors: (1) the information technology revolution, (2) strong business investment, in general, and in information processing, in particular, (3) strong growth in aggregate demand, and (4) improved managerial performance. Productivity has been boosted by the broad based nature of the computer revolution, which has increased worker productivity across different occupations and skill levels. Strong growth in aggregate demand in the late 1990's raised measured productivity by increasing the intensity that labor and capital resources are utilized as well as encouraging additional business investment Chatterjee (pp. 18-20). Relatively slow growth in the first four years of the recovery allowed the economy to avoid supply constraints in the mid 1990's. The lack of supply constraints in the mid 1990's helped fuel noninflationary growth in the second half of the 1990's when aggregate demand sharply accelerated. U.S. managerial performance has improved in response to increased global competition in goods and services markets, especially in the areas of inventory management, cost containment, and managerial control.

Measured nonfarm productivity is expected to slow to approximately 2.0 percent in 2001; its average rate over the 1960 through 1999 period. Slower nonfarm productivity in 2001 is expected primarily from slower growth in aggregate demand and slower growth in business investment spending. Underlying nonfarm labor productivity growth after removing the negative effects on measured productivity of slower growth in aggregate demand and higher inflation resulting from higher relative prices for imports, energy, and medical items should remain above 2 percent. Growth in business fixed investment spending is expected to remain strong but is expected to slow to the 5 to 7 percent range in 2001. Slower investment spending is expected to be generated by higher capital costs, a reduction in credit availability for marginal business borrowers, and a narrowing of the gap between the actual and desired capital stock.

Growth in Business Fixed Investment Spending to Remain Strong, But Slow Significantly

The boom in business fixed investment spending has accelerated in recent years. Between 1995 and 1999, business fixed investment grew at an annualized rate of 11.0 percent, up from the 7.6 percent rate between 1991 to 1995. In the first half of 2000, business fixed investment grew at a 17.7 percent annualized rate before slowing to 7.8 percent in 2000III. The stronger growth of business fixed investment since 1995 reflects the increased profitability of business investment brought about by the continuing improvements and innovation in capital goods (especially in the information technology area) and higher rates of resource utilization in general. Measuring overall expected profitability of business investment is difficult, although theoretically it is strongly related to the valuation of existing capital relative to its replacement cost (Tobin's q ratio). When the market valuation of existing capital (debt plus equity) is high relative to asset replacement cost, returns to existing capital are high and additional investment is encouraged.

Business investment spending is expected to slow to the 5 to 7 percent range due to: (1) higher capital costs, (2) reduced credit availability and increased difficulty in issuing equity securities, and (3) lower expected returns on investment in general. Figure 1 shows that Tobin's q is expected to fall through 2001 but to remain high by historical standards. External finance for marginal business borrowers in 2001 will be more expensive and difficult to obtain. In response to rising risk premiums on financial assets, equity and bond issuance fell 20 and 9 percent respectively in the third quarter relative to 1999 levels. Lower expected returns on business fixed investments, in general, is indicated by slower expected growth in corporate sales and profits, slower growth in equity and bond issuance by nonfinancial corporations thus far in 2000, and by the fall in Tobin's q ratio.

Continuation of Strong Foreign Growth with a Modest Fall in the Dollar

Strong world growth outside of the United States is expected in 2001 with growth picking up in Japan, Latin America, and Africa.. Continued strong foreign growth will produce moderate growth in the foreign demand for U.S. exports. More mature economic expansions in developed countries as well as improving financial stability in Asia should further raise foreign demand for U.S. exports, especially U.S. capital goods. Growth in western Europe is likely to be slightly lower reflecting higher European interest rates, oil prices, and inflation.

The dollar is expected to fall modestly on a broad trade weighted basis in 2001. A mild fall in the dollar is expected from the combined impact of slower growth in the U.S. relative to the rest of the world, larger trade deficits, and expected slightly lower real interest rates on government and high grade U.S. debt securities. Increased uncertainty concerning lower grade debt securities and U.S. equities will further weaken the dollar. The mild fall in the dollar will contribute to long-term adjustment in the U.S. balance of payments while avoiding destabilizing capital flight out of the United States.

Major Factors Slowing U.S. Near Term U.S. Economic Growth

Tight Labor Markets

Labor markets tightened in 1999 and the first three quarters of 2000. The unemployment rate has remained at 4.2 percent or lower since 1999III and far below empirical estimates of the NAIRU. In the last year, other measures of labor market tightness indicate a further tightening of labor markets. The percentage of those unemployed because of permanent job loss has fallen while the voluntary quit rate among the unemployed has risen. Furthermore, labor force participation reached an all time high by the summer of 2000. Tighter job markets will accelerate growth in employee compensation costs and slow employment growth in 2000 and 2001 below the 1.5 percent growth achieved in 1998 and 1999. In the first three quarters of 2000, employment growth slowed to 0.8 percent and is expected to average approximately 1.0 percent in 2001.

Current Real GDP is Above Potential GDP

Tight labor markets with the unemployment rate below all empirical estimates of the NAIRU is the major factor in actual GDP being above its estimated long run potential. Despite rapid growth in business capital and higher total factor productivity growth in this business cycle relative to business cycles of the 1970's and the 1980's, GDP remains above its long-term potential. When actual GDP exceeds potential GDP, upward pressure on inflation is normally generated as shortages of labor and capital cause production costs to rise and as firms attempt to raise profit margins. In addition, supply constraints are generated which slow growth in real GDP and act to move actual and potential GDP toward each other over time. Furthermore, rising real interest rates and tighter monetary policy is normally generated that over time reduces credit demand from interest sensitive and less financially secure borrowers. The combination of favorable relative price shocks and strong growth in aggregate demand in the 1990's allowed real economic growth to exceed its long run potential with little inflation or constraints on output. However, with the relative prices of imports and credit rising, slower growth in real output will move output closer to potential output in 2001.

Numerous private macroeconomic forecasting services and CBO produce potential GDP estimates. My empirical work expands upon previous work by Arnold in modeling potential GDP by decomposing potential GDP into its business and non business potential output components. In addition, my work decomposed potential business output into labor (labor hours worked) business capital, and total factor productivity components. The labor hour worked series consistent with nonaccelerating inflation was allowed to vary over time in response to changes in labor force growth across business cycles and in the NAIRU.[1] Total factor productivity was specified as a function of deterministic trends (which change across business cycles) and stochastic shocks. Potential business output was estimated using the Kalman filter where changes in inflation are a function of lagged changes in inflation, the differences between actual and potential business and non business output, and distributed lags of changes in the relative prices of imports, food and energy, and last quarter's growth in final goods and services prices relative to growth in unit labor costs. A further discussion of the potential GDP model and parameter estimates are presented in the Appendix.

As shown in Figure 2, business output exceeded its potential by 4.2 percent at the end of 1999. Although lack of BEA capital stock data precluded estimating the model for 2000, the gap has undoubtedly widened given the strong 4.4 percent in real GDP growth in the first three quarters of 2000, additional labor market tightening, and faster growth in unit labor costs. Figure 3, shows that nonbusiness output exceeded its estimated potential by 0.8 percent at the end of 1999. The output gap for the nonbusiness sector gap has widened in 2000 as well, given the more rapid growth in government spending in 2000. The output gaps of the business and nonbusiness sectors of the economy are likely to constrain growth more significantly in 2001.

Slower Credit Growth, Higher Interest Rates, Capital Costs, and Tighter Lending Standards

Relative abundance of funds at favorable terms in bond and equity markets financed the rapid pace of business investment that has fueled the rapid growth of the last five years. Between 1995-1999, new corporate bond and equity issuance grew at an annual rate of 15.7 percent. The 1990's also witnessed extremely rapid growth of the private (non-public) equity market, which greatly facilitated equity issuance by new venture and small existing firms (Prowse).

In 2000, rising risk premiums and liquidity concerns have slowed the issuance of corporate bonds, and equity securities. Overall new bond and equity issuance in the first three quarters of 2000 averaged 10.3 percent lower on a quarterly basis relative to 1999. Analysts are becoming increasingly concerned over much tighter conditions in bond and equity markets and the significant negative impact on economic growth in 2001 that continued tight conditions in capital markets will have on economic growth (Lonski, and Pearlstein).

The important role of credit growth in influencing the pace of economic growth is well established in the

[1] My empirical work indicated the NAIRU fell to 5.5 percent by 1999IV.

economics literature (Beranke, and Beranke and Blinder, among numerous others). Strong growth in credit has fueled our strong economic growth in recent years by providing borrowers ample credit for current investment and consumption spending. From 1997 and through 1999, real nonfinancial credit expanded at rates of 5.1, 8.2, and 8.3 percent, respectively. In the first half of 2000, real nonfinancial credit growth slowed to 3.4 percent. Rising credit standards and risk premiums on bank and nonbank business credit were major factors slowing real credit growth in the first half of 2000. Expected continued tightening of business credit terms in 2001 will further moderate growth in business credit.

Growth in consumer credit is expected to moderate as well in 2001, reflecting slower growth in consumer spending on residential housing and consumer durables. The Federal Reserve's Senior Loan Officers Opinion Survey on Bank Lending Practices indicated that credit standards on consumer loans have changed relatively little in 2000. Credit standards on consumer loans are not expected to rise significantly in 2001, given current low consumer loan default rates and large gains in households wealth in the latter half of the 1990's.

In the first three quarters of 2000, Treasury bill rose approximately 100 basis points while Treasury bond rates fell 25 to 50 basis points. Over this time period, yield spreads between corporate bond rates and Treasury bond rates (of comparable maturities) have widened substantially. For example, the spread between the BAA seasoned corporate bond and the 10 year constant maturity bond rate widened from approximately 200 basis points in 1999IV to 240 basis points in 2000III and rose to 260 basis points in October 2000.

Rising yield spreads have been even more pronounced in the non-investment grade (junk) bond area with resulting much slower issuance of non-investment grade bonds. The spread between the Standard and Poor's noninvestment grade bond yield index and the 5 year Treasury bond rate widened from approximately 540 basis points in 1999IV to approximately 690 basis in 2000III. In October, the average spread rose to approximately 800 basis points. Risk premiums on noninvestment grade bonds have reached extremely high levels for a non-recessionary period. In response to rising noninvestment grade bond yields, noninvestment grade bond issuance yearly through October 2000 was down approximately 40 percent relative to the first 10 months of 1999 (Rao). Noninvestment grade bond issuance will remain subdued in 2001 because of continued high risk premiums and expected rising bond default rates (Hamilton).

Less favorable markets for raising business capital has not been confined to securities markets. Lending standards on business loans at commercial banks have tightened progressively since the beginning of the year. The Board of Governors Senior Loan Officer Opinion Survey on Bank Lending Practices for August 2000 indicated that 34 percent of domestic banks reported tighter lending standards for loans to large and middle market firms and 24 percent of domestic banks reported tighter lending standards on small business loans. This represented a significant increase in the proportion of banks tightening business lending standards since the May survey. The Survey also indicated a rising proportion of banks have widened business lending spreads above their costs of funds, especially for riskier loans. The trend of tighter lending standards and widening lending spreads is expected to continue in 2000.
Empirical work by Lown Morgan, and Rohatgi have established the significant negative impact of rising loan spreads over funds costs and credit standards on business lending at commercial banks. Their empirical work also indicated that rising credit standards on business loans are statistically significant in depressing business investment in equipment and inventories and slowing growth industrial production, holding other factors constant.

Slower Growth in Consumer and Residential Housing Spending

Real consumer spending between 1997 and 1999 grew at an annualized rate of 4.5 percent while real consumer durable purchases grew at nearly a 10 percent annualized rate Over this same period, real residential construction grew at a robust 5.5 percent annually despite only moderate growth in the number of U.S. households. The combination of a large stock of recently purchased consumer durables and residential housing coupled with higher interest rates and much slower gains in consumer wealth indicate more moderate growth in consumer and residential housing. In the three quarters of 2000, real consumer spending on durables grew at a still robust 8.3 percent while residential investment fell 2.2 percent (at annualized rates). The expected slow down in the growth of consumer and housing spending will be tempered by continued high levels of consumer confidence and the lagged effects of sharply higher consumer wealth over the1997.

Higher Oil Prices

Higher oil prices have reduced consumer discretionary income and raised energy costs for business firms, thus reducing business profits. In addition, higher oil prices have raised short-term inflationary expectations.

However, the impact of higher energy prices is much less than in the past due largely to a significant decline in energy expenditures as a proportion of GDP relative to 20 years ago. In addition, real oil prices are still below 1980's average levels.

Oil and gas currently accounts for between 2.25 to 2.50 percent of nominal GDP as compared to 6.6 percent of nominal GDP in 1981. Furthermore, real oil prices are not abnormally high by historical standard. Real oil prices (average refiners acquisition price divided by the chain weighted consumption deflator less energy items) equaled $25.71 in 2000III. This represented a rebound from the trough in real oil prices of $12.20 recorded in 1999I. In comparison, real oil prices averaged $28.94 in the 1980's and $18.66 in the 1990's. Therefore, nominal oil prices in the $30 to $35 range when examined in real terms are not unusual and will by itself only modestly slow U.S. economic growth.

Conclusion

U.S. growth is expected to moderate to approximately 3.0 percent in 2001 after growing at an expected 5.1 percent in 2000. Growth in employment and overall employee hours is expected to average 1.0 percent while productivity growth is expected to average 2.0 percent in 2001. Slower productivity growth is expected to reflect slower overall growth in aggregate demand and moderation in the rapid pace of business investment. Tighter credit standards and greater concerns over liquidity have combined to raise capital costs for business firms, especially firms with above average default and bankruptcy risk. The risk of recession is low given the outlook for still strong (but moderating) business investment growth and the low probability of a contraction in overall consumer spending. The combined positive effect on consumer spending of continued high levels of consumer confidence and the very large gains in household wealth between 1994 and1999 make the possibility of an actual contraction in consumer spending low. With solid economic recoveries underway in Latin America and Asia and moderate growth expected in Europe, U.S. exports should continue to grow at a strong pace, thus aiding a soft landing for the U.S. economy.

Bibliography

Arnold, Robert. CBO's Method for Estimating Potential Output, Memorandum, Congressional Budget Office October 1995.

Beranke Ben S. "Alternative Explanations of the Money Income Correlation," Carnegie Rochester Conference on Public Policy. Vol. 25 (1986) pp. 49-100.

Beranke S. and Allen Blinder. Credit, Money, and Aggregate Demand. American Economic Review Vol. 78 (1988) pp. 435- 439.

Brinner Roger E. "Is Inflation Dead," Federal Reserve Bank of Boston, New England Economic Review, January/February 1999, pp. 37-49.

Browne Lynn Elaine "U.S. Economic Performance: Good Fortune, Bubble or New Era?," Federal Reserve Bank of Boston, New England Economic Review, May / June 1999, pp.3-20.

Campbell Jeffrey R. And Ellen R. Rissman. "Long-run Labor Market Dynamics and Short-run Inflation." Federal Reserve Bank of Chicago,Economic Perspectives, March-April 1994, pp. 15-27.

Chatterjee Satyajit, "Productivity Growth and the American Business Cycle, Federal Reserve Bank of Philadelphia,Business Review, September / October 1995, pp.13-22.

Duca, John V. "What Credit Market Indicators Tell Us," Federal Reserve Bank of Dallas,Economic and Financial Review Q3 1999, pp. 2-13.

Federal Reserve Board of Governors. Washington D.C. Senior Loan Officer Opinion Survey on Bank Lending Practices, November 2000.

Hamilton David T. "Third Quarter Default Commentary: Default Pace Slows in the Third Quarter, Poised for Increase," Moody's Investors Service New York N.Y. Default Research October 2000.

Kuttner Kenneth N. Using Noisy Indicators to Measure Potential Output Federal Reserve Bank of Chicago, Working Paper 91-14, June 1991.

Kuttner Kenneth N. "Monetary Policy with Uncertain Estimates of Potential Output." Federal Reserve Bank of Chicago, Economic Perspectives, January-February 1992, pp.2-15.

Kuttner Kenneth N. "Estimating Potential Output as a Latent Variable," Journal of Business and Statistics, Vo. 12, No.3 (July) 1994. pp. 361-368.

Lonski John. "Private Sector Borrowing Restraint May Slow GDP," Moody's Investors Service New York N.Y. Economic Commentary, November 6, 2000.

Lown Cara S., Donald P. Morgan, and Sonali Rohatgi. "Listening to Loan Officers: The Impact of Commercial Credit Standards on Lending and Output," Federal Reserve Bank of New York,Economic Policy Review, pp.1-16.

Pearlstein Stephen. "Economy Brakes as Funds Slow: Big Investors Rethink New Firm Stakes. Washington Post, October 30 2000, p.1.

Prowse, Stephen. "The Economics of the Private Equity Market," Federal Reserve Bank of Dallas,Economic Review, Third Quarter 1998, pp. 21-34.

Rao Kamelesh. "Junk Bond Yields Jump Higher," Moody's Investors Service Economic Commentary, November 6, 2000.

Rich Robert W. and Donald Rissmiller. "Understanding the Recent Behavior of Inflation,"Federal Reserve Bank of New York, Current Issues in Economics and Finance, July 2000.

Van Horne, James C. "The Default Risk Structure of Interest Rates," (Chapter 7) inFinancial Markets Rates and Flows (Fourth Edition), Prentice Hall, Englewood Cliffs, New Jersey 1994.

Appendix

Potential business and nonbusiness output are estimated as not directly observable variables that enter the inflation generating process. Specifically, in the absence of relative price shocks, higher inflation is expected when business and nonbusiness output exceed their long-term potential levels. Potential business output is specified in a production function format and is estimated directly using the Kalman filter. The Kalman filter estimates the values of potential business output over time that (given the specified structures for potential nonbusiness output and the change in inflation equations) maximizes the log likelihood function for change in inflation equation. Due to difficulties in estimating output and factor input usage for the nonbusiness sector, nonbusiness sector potential output is modeled outside of the model using a modified segmented trend business cycle approach. This approach is similar to the BEA's approach in estimating potential nonbusiness output (Arnold, pp.11-16).

The model expands upon the work of Kuttner(1991, 1992, and 1994) allowing for variability across business cycles in the deterministic component of total factor productivity and by examining the relative importance of business and nonbusiness output gaps in the inflation generating process. The business sector is much larger accounting for nearly 86 of total GDP in 2000II. Given the much greater relative size of the business sector, accurate measurement of potential business output is critical in measuring potential GDP. Furthermore, the business sector is likely to be more sensitive in pricing its output in response to changes in overall aggregate demand conditions, given the great importance of profit and market share goals to business organizations. The impact of overall goods and services prices to excess demand or supply in the nonbusiness sector is likely to be more muted, especially in the short and intermediate term, and is more likely to reflect longer term resource costs considerations.

Empirical results indicated that the business output gap was significant at the one percent in explaining changes in inflation while the nonbusiness output gap term was insignificant in explaining changes in inflation. Business output that is one percent above potential was found to generate an approximate .07 percent increase in inflation in the current quarter.

Model Specification

The potential GDP model is specified below:

(1) $\triangle \mathrm{Inf}_t = \sum B_i \triangle \mathrm{Inf}_{t-1} + \sum C_i \triangle \mathrm{rimp}_{t-I} + \sum D_i \mathrm{rfe}_{t-I} + \sum E\ \mathrm{devinfulab}_{t-1} + F(\mathrm{busout} - \mathrm{potbusout})_t + G\ (\mathrm{nonbusout} - \mathrm{potnonbusout})_t + U1_t$

where
inf= Inflation (chain weighted GDP deflator)

rimp = Real import prices (chain wt. import price deflator / chain wt. GDP deflator)

Rfe = Real food and energy prices (chain wt. consumption deflator / chain wt. consumption deflator without food and energy)

devinfulab = inflation (chain weighted GDP deflator) minus inflation in unit labor costs

potbusout = log of potential business output in the absence of relative price shocks.

(2) $\mathrm{potbusout} = (.7)\triangle \mathrm{Noninflabhrs}_t + (.3)\triangle \mathrm{Capstk}_t + D\ D741_t + E\ D802 + F\ D814 + G\ D903 + \mathrm{potbusout}_{t-1} + U2_t$

busout = log business output (chain wt.)

nonbusout = log of the sum of government, household, nonprofit institutions output and the GDP sector output residual.

potnonbusout = log of nonaccelerating inflation nonbusiness sector output. Generated from regression of nonbusiness output on a constant, business cycle trend dummy variables, an early 1980's Reagan government spending slowdown dummy variable (DREAGAN), and the gap between the unemployment rate and the NAIRU.

After estimating regression, the gap between unemployment rate and the NAIRU was set equal to zero to remove business cycle influences and the fitted values from regression were set equal to potnonbusout. Similar approach to cyclical adjustment procedure used by CBO (Arnold pp. 11-13).

noninflabhrs = noninflationary business labor hours. Actual business labor hours are adjusted to a level consistent with estimated NAIRU. Constructed from a regression of the logarithm of actual labor hours on a constant, trend business cycle variables and the gap between the unemployment rate and the NAIRU. After estimating the equation, the unemployment rate NAIRU gap is set equal to zero to remove business cycle influences on business labor hours. Fitted values from the NAIRU gap adjusted equation are set equal to noninflabhrs. Similar to cyclical adjustment procedure used by CBO (Arnold pp. 7-8).

NAIRU = nonaccelerating inflation rate of unemployment. Time varying NAIRU estimated by substituting, the gap between the unemployment rate and the NAIRU for the business and nonbusiness output gap terms in the change in inflation equation and estimating using the Kalman filter. Estimated NAIRU for 1999IV was 5.5 percent

$U1_t$ = stochastic first order autoregressive error term of the form $U1_t = p\ U1_t + e_t$

$U2_t$ = stochastic error term. Best results obtained with variance of .000001.

Business Cycle Trend Dummy Variables:

D741 = business cycle variable to measure trend productivity for 74I to 80I1 business cycle. Takes value of one for the 74I to 80I period and zero outside of period.

D802 = business cycle variable to measure trend productivity in the 80II-80III and 804IV 81III business cycles. Takes a value of one for the 80II to 81III period and zero outside of period

D814 = business cycle variable to measure trend productivity in the 81IV to 90II business cycle Takes a value of 1 for the 814 to 813 period and zero outside of period

D903 = business cycle variable to measure trend productivity in the 90III814 to 99IV period. Takes a value of 1 for the 903 through 994 and zero outside of the period

DREAGAN = Dummy variable to capture the slowdown in government spending over the 1980III to 1983IV period.

Table 1
Change in Inflation
Nonbusiness Output Gap Included

	Coefficient	Std. Error	z-Statistic	Prob.
Change $Inf_{(t-1)}$	0.247027	0.078421	3.149998	0.0016
Change $Inf_{(t-2)}$	-0.425431	0.083324	-5.105754	0.0000
Change $Rimp_{(t)}$	-0.037679	0.012738	-2.957988	0.0031
Change $Rimp_{(t-1)}$	0.084600	0.017676	4.786021	0.0000
Change $Rimp_{(t-2)}$	-0.018102	0.015724	-1.151263	0.2496
Change $Rfe_{(t)}$	0.500957	0.101918	4.915274	0.0000
Change $Rfe_{(t-1)}$	-0.690934	0.153574	-4.499027	0.0000
Change $Rfe_{(t-2)}$	0.170809	0.134172	1.273062	0.2030
$devinfulab_{(t-1)}$	-0.049993	0.025020	-1.998127	0.0457
$(busout\ -potbusout)_{(t)}$	7.543822	2.881355	2.618151	0.0088
$(nonbusout-potbusout)_t$	-3.217646	17.82164	-0.180547	0.8567
$U1_{(t-1)}$	-0.574231	0.110106	-5.215255	0.0000
var $e_{(t)}$	-0.598055	0.154536	-3.869996	0.0001
D741	0.003535	0.001191	2.966992	0.0030
D802	-0.004776	0.006959	-0.686341	0.4925
D814	0.003449	0.001159	2.975765	0.0029
D903	0.004000	0.001088	3.676758	0.0002
	Final State	**Root MSE**	**z-Statistic**	**Prob.**
logpotbusout (1999IV)	8.925441	0.007809	1142.922	0.0000

Log likelihood	-118.3898
Parameters	17
Akaike info criterion	2.603650
Schwarz criterion	3.035906

Table 2
Change in Inflation
Nonbusiness Output Gap Excluded

	Coefficient	Std. Error	z-Statistic	Prob.
Change $inf_{(t-1)}$	0.224894	0.078064	2.880901	0.0040
Change $inf_{(t-2)}$	-0.407815	0.081613	-4.992104	0.0000
Change $rimp_{(t)}$	-0.037815	0.012972	-2.915180	0.0036
Change $rimp_{(t-1)}$	0.083722	0.017868	4.685652	0.0000
Change $rimp_{(t-2)}$	-0.016037	0.015432	-1.039208	0.2987
Change $rfe_{(t)}$	0.498445	0.101673	4.902436	0.0000
Change $rfe_{(t-1)}$	-0.673297	0.151561	-4.442416	0.0000
Change $rfe_{(t-2)}$	0.149449	0.125753	1.188437	0.2347
$devinfulab_{(t)}$	-0.050302	0.024214	-2.077353	0.0378
$(busout - potbusout)_{(t)}$	7.210295	2.717842	2.652949	0.0080
$U1_{(t-1)}$	-0.558310	0.115032	-4.853518	0.0000
var $e_{(t)}$	-0.611063	0.152201	-4.014835	0.0001
D741	0.003464	0.001338	2.797161	0.0052
D802	-0.004201	0.007214	-0.582404	0.5603
D814	0.003387	0.001221	2.773506	0.0055
D903	0.003995	0.001133	3.525170	0.0004
	Final State	Root MSE	z-Statistic	Prob.
logpotbusout (1999IV)	8.924873	0.007968	1120.092	0.0000

Log likelihood	-118.5541
Parameters	16
Akaike info criterion	2.587580
Schwarz criterion	2.994409

Figure 1. Tobin's Q* Peaked in 2000QI and Will Continue to Fall in 2001

*Tobin's q is the market value of debt and equity
divided by the the replacement value of assets less non debt liabilities of nonfinancial corporations

Figure 2. Actual Business Output is Above Potential Output

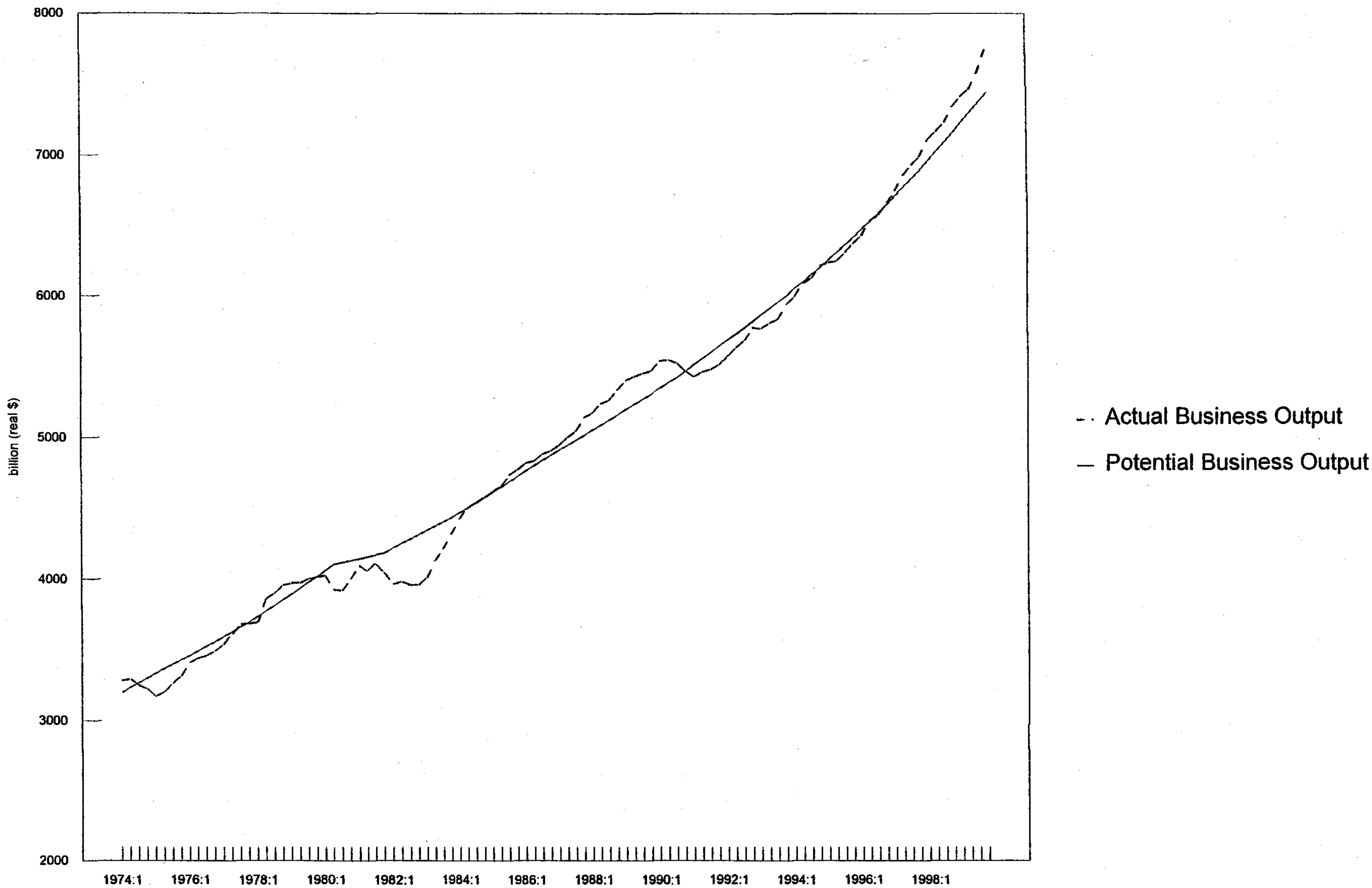
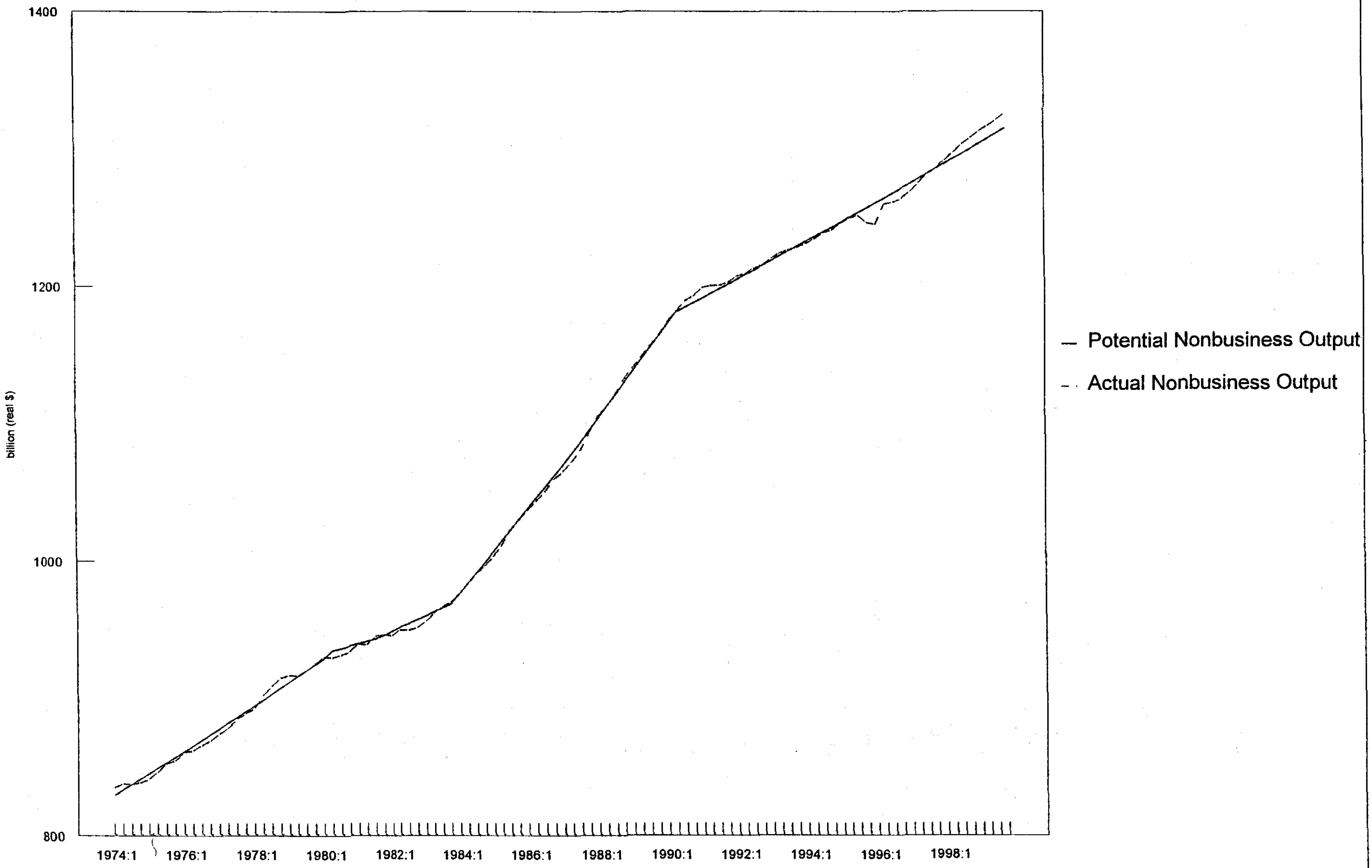


Figure 3. Nonbusiness Output Is Slightly Above Potential Output



The Outlook for Productivity Growth: Are We in a New Economy?

Robert W. Arnold
Macroeconomic Division
Congressional Budget Office

Whether you believe that you are in a "New Economy" depends on what you mean by the term–it means different things to different people. I'll take it to describe the extraordinary confluence of good economic news during the second half of the 1990s. This battery of good news includes rapid economic growth, a falling unemployment rate, robust productivity growth, declining inflation, and very strong business investment, with a decided tilt toward information technology (IT) goods.

Twice each year, CBO assembles an economic forecast and projection that is used as input to the agency's budget projections. Since CBO's mandate is to produce nonpartisan analysis, the economic forecast is meant to reflect a consensus of private and government forecasters. The current forecast was released in July 2000. In it, CBO projects that growth in real GDP will moderate to a 2.7 percent rate during the 2000-2010 period, while inflation (measured by the CPI-U) averages 2.6 percent, and labor productivity grows at a 2.2 percent pace. I will talk about how CBO arrived at those projections and discuss whether they reflect a New Economy.

Fundamentally, CBO projects that real growth will slow because, in the agency's view, the economy is operating at a high degree of resource use and is straining its productive capacity. This judgement is based, in part, on the output gap, or difference in percent between real GDP and potential GDP (see Figure 1). Potential GDP, defined as the level of real GDP that is consistent with stable inflation, is estimated using the nonaccelerating inflation rate of unemployment (or NAIRU) as a benchmark. Based on past patterns, a positive output gap suggests that growth will slow and that inflation will rise. Note that potential growth remains fairly rapid in CBO's projection, averaging 3.1 percent between 2000 and 2010.

A Digression on the NAIRU

The NAIRU has received a lot of criticism lately–low unemployment combined with falling inflation will cause that. Some argue that the NAIRU is lower than CBO's estimate, which is 5.2 percent currently. Some argue that the concept should be scrapped altogether. Although CBO has deemphasized the NAIRU in our thinking, it hasn't jettisoned it completely. The basic story provided by the measure–that the labor market is tight–is confirmed by independent evidence. Indeed, it would be hard to believe that the NAIRU has permanently fallen as low as 4 percent.

One reason not to abandon the concept of the NAIRU (and the underlying Phillips curve that is used to estimate it) is that there has been a good correlation over time between the unemployment gap and changes in inflation (see Figure 2). It's not perfect, but no statistical relationship is. Perhaps the NAIRU is a victim of its success–it worked very well during the late 1980s, and that might have fostered unrealistic expectations about its forecasting ability. People forget that the unemployment gap is but one of many factors influencing inflation. However, the most important reason not to abandon the NAIRU is that, unlike price-based Phillips curves, wage-based Phillips curves are still tracking the data reasonably well. Measures of wages and, to a lesser degree, compensation are behaving about the way the theory would predict (see Figure 3). For this reason, CBO believes that no fundamental change has occurred in the way labor markets work.

NOTE: Robert Arnold is a Principal Analyst in the Macroeconomic Analysis Division of the Congressional Budget Office. Although this paper draws on publications by the CBO, the views expressed are those of the author and should not be interpreted as those of CBO.

The real puzzle about the late 1990s is why price inflation remained muted in the face of tight resources and accelerating growth in compensation. CBO looked for factors holding down price inflation, but not wage inflation, and found several that could be interpreted as beneficial supply shocks:

- o Import prices. The price of imported goods fell dramatically during the late 1990s. By itself, that decline could have knocked a full percentage point off the inflation rate during the 1996-1999 period.

- o Computer prices. The price index for computers and peripherals has been declining for as long as the data exist, but the rate of decline accelerated dramatically during the late 1990s. I'll have more to say on this later.

- o Measurement changes. The Bureau of Labor Statistics introduced a host of methodological changes in recent years that reduced the measured rate of inflation. Of course, when BLS changes its inflation formula, nominal spending does not change. So when BLS revised down the inflation rate, real growth was revised up one-for-one. This had a small impact, accounting for about 0.1percentage point of the missing inflation.

<u>The Key: Productivity Growth</u>

The most important factor restraining price inflation is faster labor productivity growth–faster growth in wages and compensation does not have to feed through to prices if productivity growth increases also. Inflation is closely correlated with unit labor cost, which is defined as compensation per unit of output, and is calculated as compensation per hour divided by output per hour (see Figure 4). The recent increase in compensation growth does not show up in prices because faster productivity growth held down unit labor cost. Productivity had been growing along a fairly constant trend of about 1.4 percent since 1973. During the last four years, however, it surged to a rate of 2.7 percent. And for the year ending in the second quarter of 2000, it spiked to a 4.5 percent rate of growth (see Figures 5 and 6).

When we project labor productivity, what should we do with the last four years of data? Should we continue that recent trend? Should we ignore it and go back to the 1973-1995 trend? Our task is complicated by the fact that productivity growth–including the source of the post-1973 slowdown–is not well understood. If the economy is operating above its potential, then some of the productivity surge is cyclical and will therefore reverse itself. Moreover, five years of data is not enough to reliably estimate a trend. However, it could be early evidence of a return to the glory days of the 1950s and 1960s.

CBO searched for factors underlying the upswing in productivity growth to determine if they would persist. CBO came up with three factors, two of which relate to the discussion about a New Economy and one which does not. I'll discuss the easiest one first.

- o Measurement Changes. The measurement changes described earlier that reduced measured inflation raised the measured growth rate and the rate of productivity growth. By CBO's estimate, those changes contributed about 0.1 percentage point to the productivity acceleration–and they had nothing to do with the New Economy.

- o Capital Deepening. One important feature of the recent economic picture is the boom in business fixed investment. The neoclassical model of long-run growth implies that the amount of capital per worker will correlate with productivity growth. That correlation reflects the effect of capital goods, including computers and other IT capital, being used to produce other goods. The correlation is hard to see in the year-to-year changes in productivity and capital per worker. However, it is easier to see once the data have been smoothed (see Figure 7). Part of the reason why labor productivity growth accelerated is because investment has been so strong. By CBO's estimate, capital deepening accounts for 0.4 percentage point of the 1.1 percentage point surge in labor productivity.

- The Quality of Computers. Along with Macro Advisers, Oliner & Sichel, Kevin Stiroh, and Robert Gordon, CBO attributes some of the upswing in productivity growth to faster technical change in the production of computers. Even casual observers are aware that computers have become steadily more capable through the years, with dramatic increases in speed and storage capacity. Those quality improvements represent higher productivity in the production of computers (and their components) and show up in the data as falling computer prices.

  Apparently, there was a major shift upward in the productivity of the computer sector during the late 1990s. The price of computers, which had been declining at a rate of roughly 13 percent since the early 1970s, started plunging at nearly 30 percent starting in 1996. It is possible to estimate the contribution of the decline in computer prices to productivity growth using the method pioneered by Kevin Stiroh and the other researchers. By CBO's estimate, faster productivity growth in the computer sector accounts for 0.2 percentage points of the overall productivity acceleration, about the same as the estimates made by the other researchers.

Even after accounting for those factors, CBO found additional growth in total factor productivity, so the agency boosted the projected growth rate by 0.1 percentage point to reflect the possibility that there is a faster trend in productivity growth.

CBO's Projection

The effects of each of those factors are outlined in Figure 8, which shows the projection for potential labor productivity from the agency's July 2000 forecast. Had CBO merely continued the trend that labor productivity has been following since 1973, the agency would project it to grow at a 1.5 percent rate during the 2000-2010 period. Adding the effects of capital deepening (which depend on CBO's forecast for business investment) raises the projected growth rate by 0.4 percentage point. The other factors discussed above collectively add another 0.4, raising the projected growth rate to 2.3 percent.

Is this a New Economy projection? Yes and no. It is a New Economy projection in the sense that it includes most of the post-1995 acceleration in labor productivity. However, it is not a New Economy projection in the sense that investments in IT are boosting productivity growth elsewhere in the economy. Also, not all of the factors that CBO has identified will persist indefinitely. For example, the increased pace of capital deepening will taper off when the investment boom ends.

Conclusion

CBO has produced a mildly optimistic view of the outlook for the new economy. One lesson: the amount of uncertainty associated with medium-term projections, which is always high, is even greater than usual.

Figure 1: The Output Gap

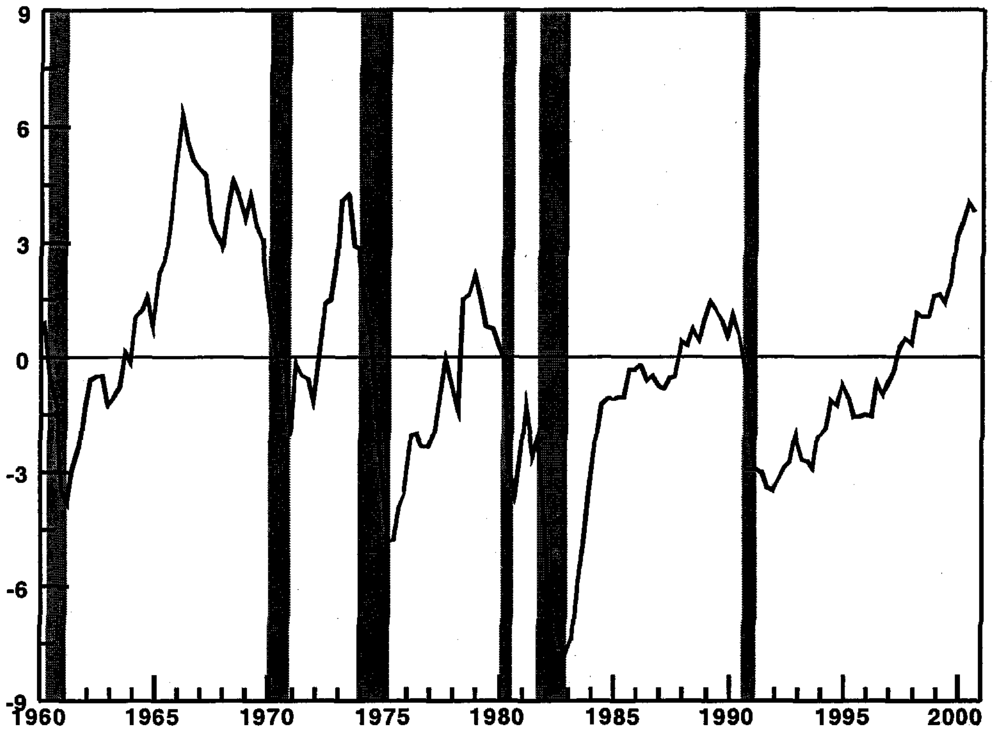
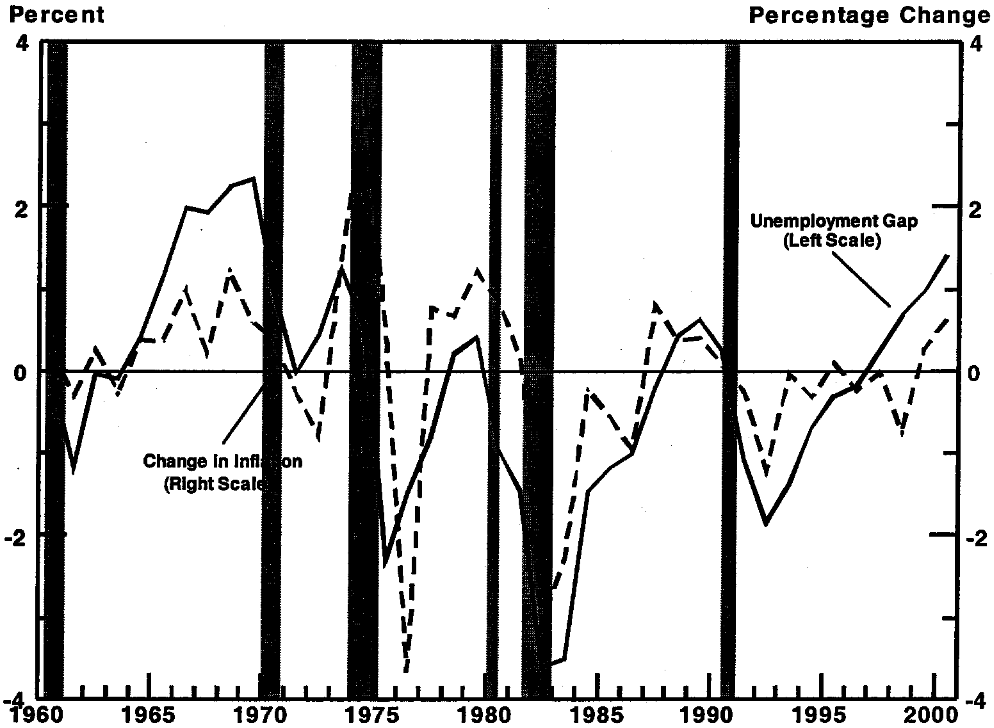


Figure 2: Labor Market Tightness and the Change in Inflation



Notes: Inflation measured using the GDP price index.

The unemployment gap is defined as the difference between the NAIRU and the civilian unemployment rate.

Figure 3: Employment Cost Index

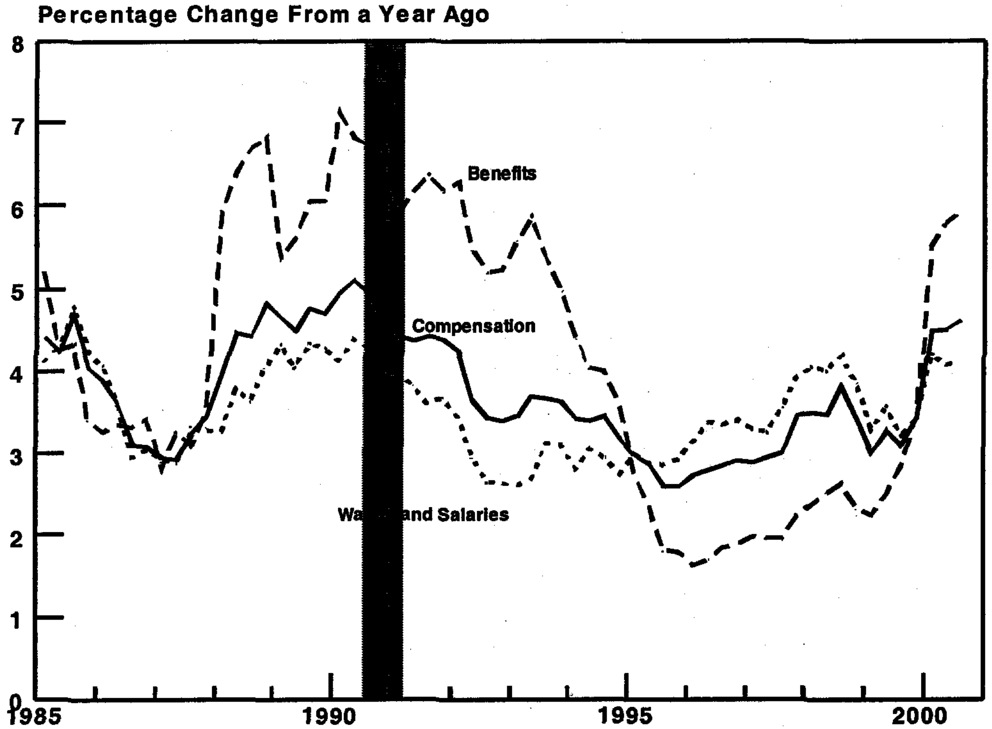
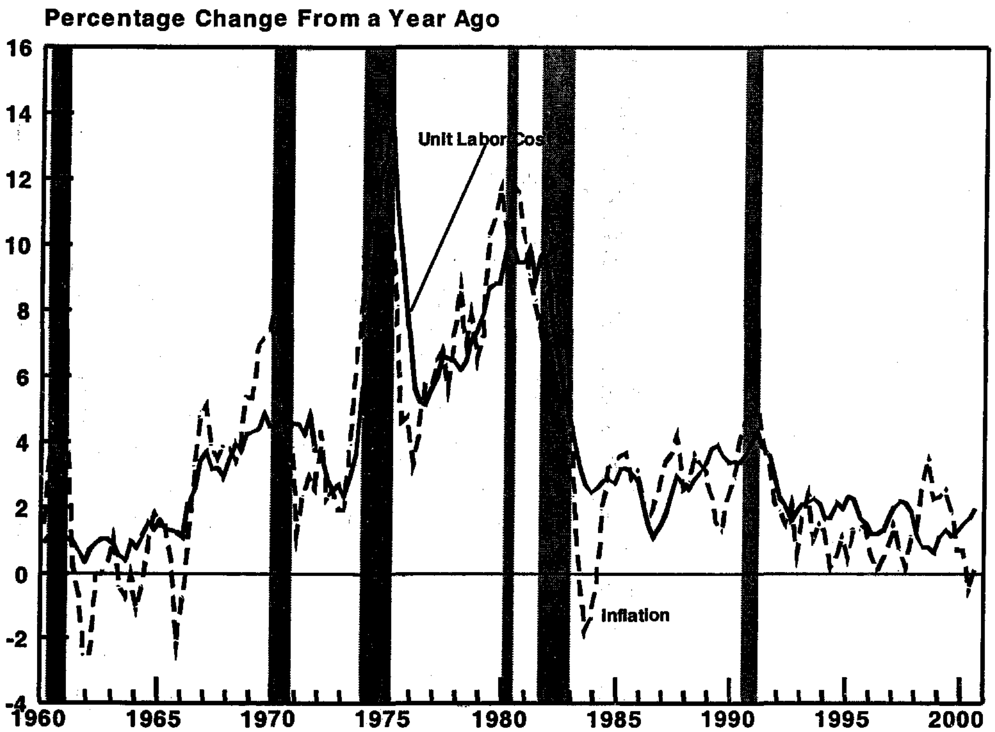


Figure 4: Inflation and Unit Labor Costs



Note: Inflation measured using the price index for GDP in the nonfarm business sector.

Figure 5: Labor Productivity and Trend

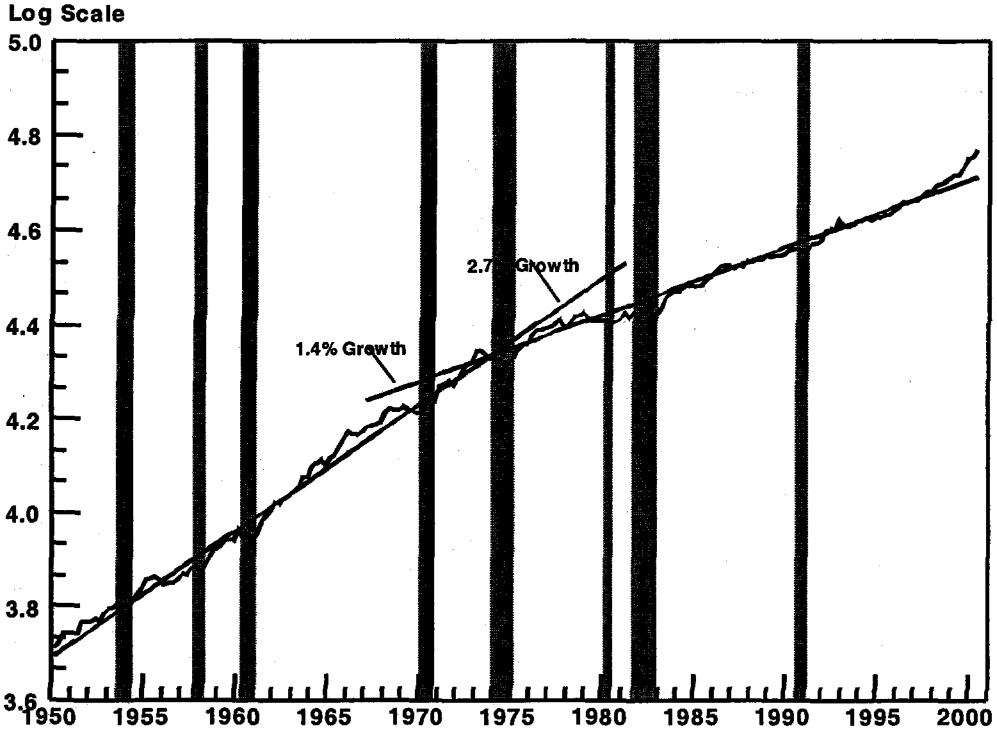


Figure 6: Growth in Labor Productivity

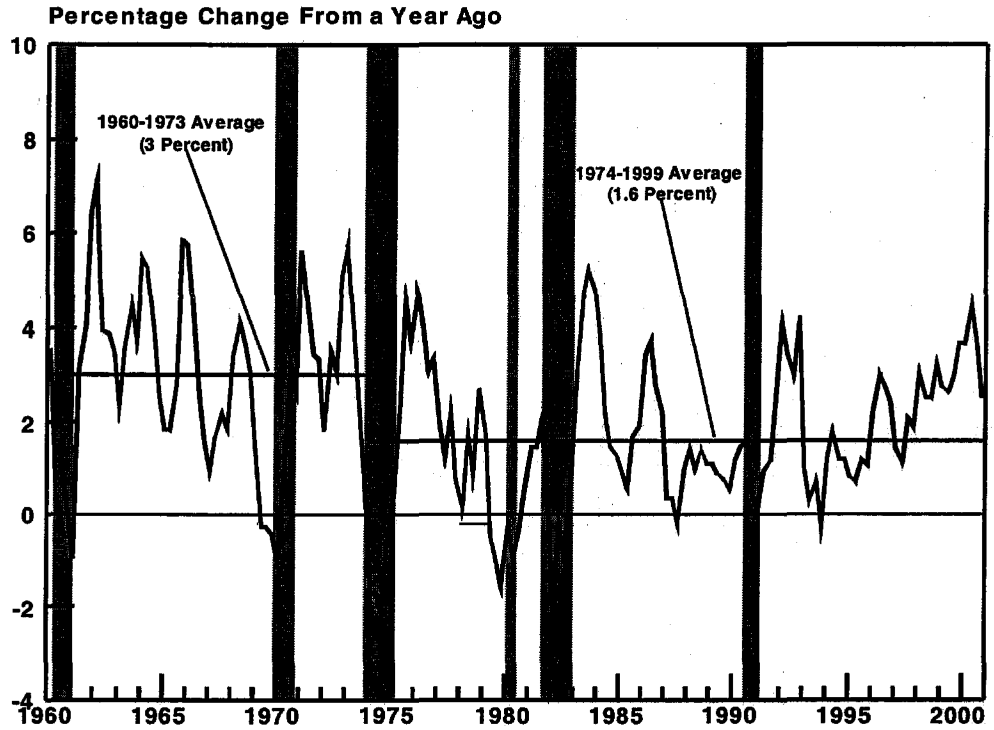
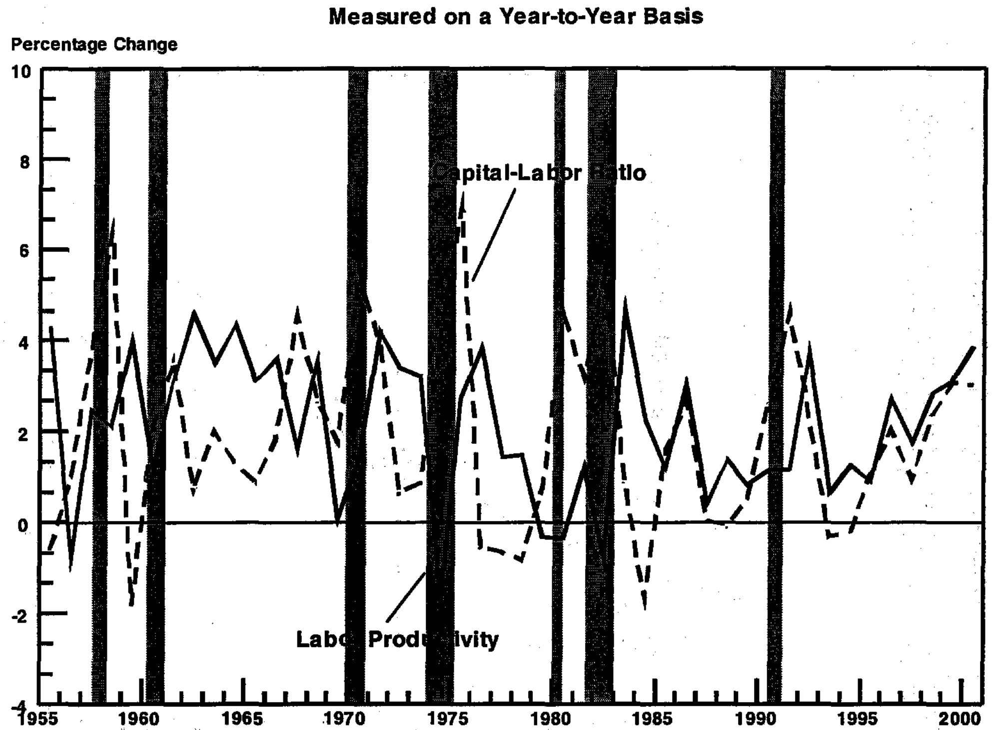


Figure 7: Growth in Labor Productivity and the Capital-Labor Ratio



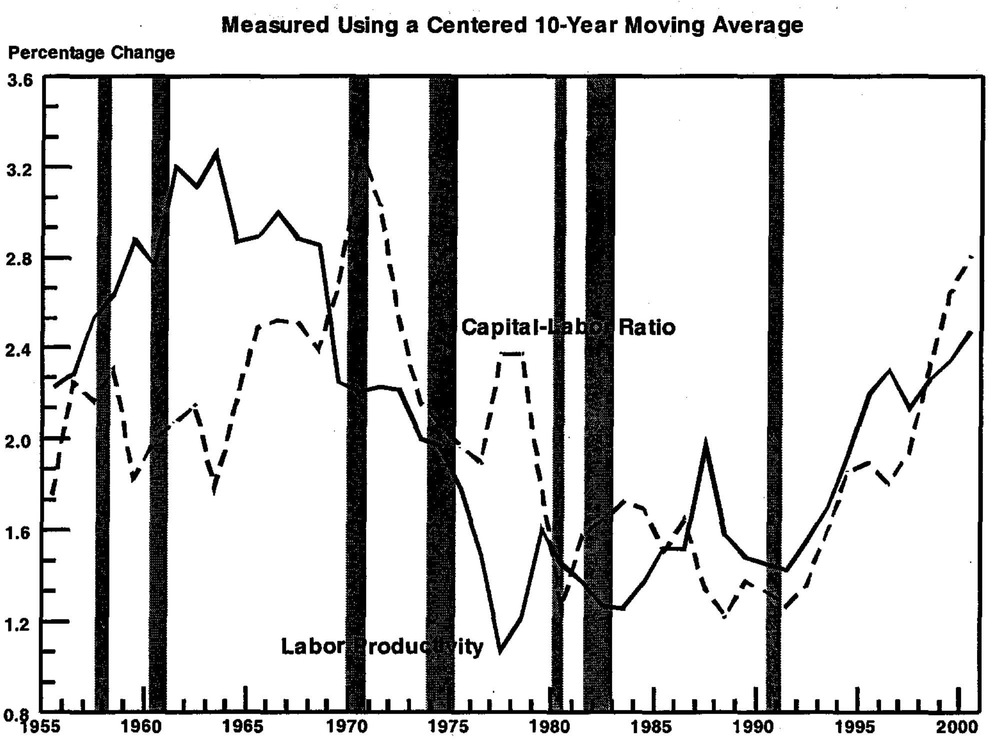
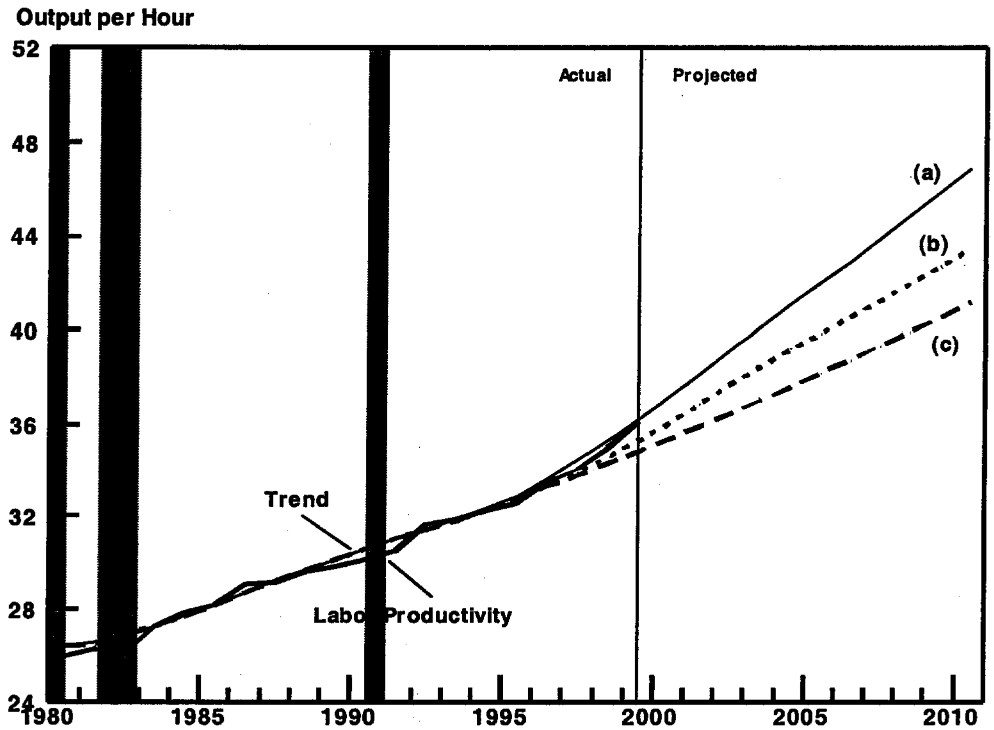


Figure 8: Factors Affecting Labor Productivity



(a) Trend growth plus capital deepening plus the effect of adjustments for measurement changes, computer quality, and the possibility of faster trend.

(b) Trend growth plus the effect of capital deepening.

(c) Trend growth in labor productivity.

HEALTH SECTOR FORECASTING AND POLICY ISSUES

Chair: Herbert Traxler
Bureau of Health Professions
U. S. Department of Health and Human Services

Forecasting Prescription Drug Utilization, Including the Impact of Medicare Expansion—Abstract,
Walter Bottiny and Jim Cultice, Bureau of Health Professions
U.S. Department of Health and Human Services

The Dental Requirements Model (DRM): Forecasting the Dentist Requirements for Low-Income Children,
Judith A. Cooksey, MD, MPH
Gayle R. Byck, PHD, University of Illinois at Chicago

Forecasting the Physician Workforce,
Richard A. Cooper, MD, Medical College of Wisconsin

Forecasting Prescription Drug Utilization, Including the Impact of Medicare Expansion

Walter Bottiny and James Cultice
Bureau of Health Professions, U.S. Department of Health and Human Services

As mandated by Congress, the Bureau of Health Professions is currently conducting a study to determine whether, and to what extent there is a shortage of licensed pharmacists. BHPr assessment of the extent of the pharmacist shortage will include a projection of prescription volume in the year 2005. That projection will be generated from a model that accounts explicitly for changes in demographic and health status variables, and will account for increased prescription utilization due to Medicare prescription drug coverage. This presentation will discuss the prescription volume projections and how they were derived.

THE DENTAL REQUIREMENTS MODEL (DRM): FORECASTING THE DENTIST REQUIREMENTS FOR LOW-INCOME CHILDREN

Judith A. Cooksey, MD, MPH and Gayle R. Byck, PhD
Illinois Center for Health Workforces Studies, University of Illinois at Chicago

Introduction
The Dental Requirements Model (DRM) was developed by Vector Research Incorporated (VRI) in 1999 under contract with the Bureau of Health Professions of the Health Resources and Services Administration (HRSA). The model estimates the dentist requirements to provide care to children with coverage by the State Children's Health Insurance Program (SCHIP or CHIP), a federal-state health insurance program authorized by Congress in 1997 as Title XXI of the Social Security Act. A dental need-based model was developed since the demand or utilization of dental care by low-income children has been far below the desired levels.

This paper will review the policy purpose of the model, then provide background on the dental workforce, the eligible CHIP population, and children's dental health needs. The model will be presented with indications of the model assumptions, the user inputs, and the model outputs. A state-level application of the model will be presented and discussed using Illinois data. Potential model enhancements will be discussed.

Analytic and Policy Purpose of the Dental Requirements Model
Children's dental health has improved over the past forty years, due to fluoridation, improved oral and dental hygiene, better nutrition, and access to dental care. Although dental caries rates (decayed teeth or cavities) have declined, the most recent national population survey document the continued presence of caries and substantial variation in the numbers of decayed teeth – with higher rates among older children, ethnic and racial minorities, and low-income children ((Brown, et al., 1999; Vargas, et al., 1998; Edelstein, 1995). Thus, there is still significant need for children for children to receive both preventive and restorative dental care.

The Oral Health Initiative (OHI), a joint project of HRSA and the Health Care Financing Administration (HCFA) has identified dental caries as one of the most common childhood health problems which is progressive and not self limited (US DHHS, 2000). About 25% of children (principally low-income) have untreated caries, and these children have about 80% of the population estimates of untreated caries in permanent teeth – a significant health disparity. (GAO, 2000; US DHHS, 2000). Two federal agencies, HRSA, with a commitment to access to care, and HCFA, the administrator of Medicaid, have placed a high priority on improving the dental health status of children at risk and a key strategy is to increase access to dental care. The recent Surgeon General's report on oral health noted the "silent epidemic" of dental disease and the importance of oral and dental health to general health status (US DHHS, 2000).

Based on the national 1996 Medical Expenditures Panel Survey, overall 43% of children had at least one visit to a dentist, with an estimated 87 million total visits, or 2.7 visits per child using services (Edelstein, 2000). Only about 25% of children with Medicaid have visited a dentist in a year (US DHHS, 2000). A key factor limiting Medicaid children's access has been the low participation rates of dentists, that is few dentists accept Medicaid covered children in their practices. Studies have identified low Medicaid payment rates as the most important barrier for dentists followed by billing and administrative burdens and poor patient compliance with keeping appointments (Center for Research and Public Policy, 1999, Nainar, 1996; Venezie, 1993). Dental practice in the US is a private practice model and most dentists have very limited abilities to cost shift for patients who cannot pay the costs of care.

When fully implemented, the CHIP program will bring health insurance coverage to over seven million children at or below 200% of the poverty level and in 48 states this will include dental coverage. During 1999, two million children had been enrolled in CHIP (Smith, 2000). With the low utilization rates of dental care among children with Medicaid coverage, there has been concern about access to dental care services for the CHIP-enrolled children. Particularly, since many states have implemented CHIP through a Medicaid expansion and/or are using the same provider networks for CHIP and Medicaid children. This concern led HRSA to commission the development of the DRM to assist state and federal health policy groups in planning for dental care needs for CHIP eligible children. While the model was developed for estimating dentist requirements for children with CHIP, it can also be used to estimate dentist needs for children with Medicaid.

Dental Workforce

Over the last twenty years, there has been a modest growth in the dental workforce supply in the U.S. with the count of *active* dentists increasing from 121,900 in 1980, to 147,500 in 1990, and 154,900 in 1996 (US DHHS,HRSA, 1999). However, in the 1990s, the increase in the number of dentists fell below the overall population growth. Thus the ratio of dentists to 100,000 went from 53.2 dentists per 100,000 population in 1980, to 58.7 in 1990, and to 58.1 in 1996. This ratio is projected to further decline to 56 in 2000 to 55 in 2010. (ADA 1999) This constriction of the supply of dentist is expected to have a continuing negative effect on access to care for low-income and other underserved population groups.

The most detailed data on the dental workforce comes from the American Dental Association (ADA) which conducts surveys including a census of all known dentists in the U.S. (ADA members and non-members), and annual surveys of dental practice (ADA 2000). Of the estimated 183,000 dentists in the U.S. in 1997, 149,350 were professionally active, with the remainder retired, otherwise not working in dentistry, or with missing practice data (Table 1).

Of the dentists who were professionally active, almost 93% were in private practice, others were dental school faculty, employed by the armed services, other federal, state, and local government employees, in other health organizations, or in graduate dental training. Eighty-one percent of dentists practice as general dentists with the remaining classified in the eight specialties of dentistry.

A relevant point for the DRM model is the large number of general dentists and the relatively small number of pediatric dentists, about 2.3% of dentists in private practice. The model allows users to indicate the estimated volume of dental care provided by general dentists and pediatric dentists by age group of children.

Table 1 ADA Census of Dentists by Professional Activity, 1997

Dentist Category	No. of Dentists
All dentists	183,000
Professionally Active	149,350
Private Practice Dentists	138,449
Private practice dentists in:	
General practice	112,190
Orthodontics	8,095
Oral & Maxillofacial Surg	5,179
Pediatric dentistry	3,305
Other specialties*	9,680

* Includes endodontics, periodontics, prosthodontics, oral and maxillofacial pathology, and public health dentistry.

Pediatric dentists are considered a specialty of dentistry with training in the management of children with complex medical and dental conditions and psychosocial needs, including children with disabilities. Pediatric dentists also provide dental care to healthy children, and they often locate their practices in metropolitan and suburban areas. There is no source of data on the portion of children's dental care that is provided by the general dentist versus the pediatric dentist. The major provider of dental care for children at the national level is expected to be general dentists. The model uses a pediatric dental care default value of 100% for children under 3 years of age and only 6.6 % for other ages. This estimate is based on expert opinion and can be varied by the user input.

The ADA surveys of dentists in private practice provide a useful source of information on the work patterns of dentists and their productivity (ADA, 2000). The 1998 survey reported for dentists in private practice an average of 47.6 weeks worked per year, with 36.9 hours per week in the office, and 33.4 hours per week in direct patient care. Dentists' productivity, in terms of visits per year, varies substantially with the use of dental hygienists, with 2,640 visits per year for dentists without hygienists and 3,740 for dentists with hygienists. Currently the model assumes a 2,000 hour work week although this can be varied and planned model enhancements will probably reflect a value closer to the ADA estimates.

The Population Eligible for the Child Health Insurance Program (CHIP)

The CHIP program was passed by Congress as a joint federal-state program that required each state to indicate whether the state would implement a Medicaid expansion, a new program, or a combination of the two options. It was intended to provide health (and dental) insurance to low-income families who earned too much to qualify for Medicaid but could not afford to purchase insurance coverage for their children. Implementation of CHIP has been slow, with about two million children enrolled at some point during fiscal year 1999 (Smith, 2000).

A study using national survey data found that children who were uninsured and CHIP-eligible were different from Medicaid-enrolled and privately insured groups of children in terms of sociodemographic, family-level, and health status characteristics (Byck 2000). This study found that relative to the Medicaid-enrolled population, the CHIP population is proportionately older, less minority, more likely to live in suburban and rural areas, and live in better educated and more two-parent families; they are also in better health and have fewer chronic health conditions and activity limitations. When compared to the privately insured group, the CHIP group is more likely to be Hispanic, live in urban areas, and also live in households with parents/guardians who are less educated and less likely to both be employed, as well as in fewer two-parent families. With regard to dental health care needs, CHIP children were significantly more likely than Medicaid-enrolled children and privately insured children to experience a delay or unmet dental need.

Variations in Children's Dental Care Utilization and Needs for Care

Routine dental care for children includes screening exams, preventive services (such as applications of fluoride and sealants), and restorative care (such as filling decayed teeth). A standard measure of dental care needs is the assessment of the number of teeth (T) or tooth surfaces (S) that are decayed (D), missing (M) or filled (F). A dental exam can provide a DMFT score, for the number of permanent teeth that are decayed, filled , or missing. The percent of teeth that are decayed, and not filled, indicates the need for restorative care.

National population surveys have noted a decline over time in overall children's population DMFT scores, with a consistently higher score among older children. For example, in 1963-70, the DFMT score for children 6 to 11 years of age was 1.4 with 36% of the teeth being decayed, this declined to 0.6 with 25% decayed teeth in 1988-94. Among adolescents aged 12 to 17 years, the DFMT score was 6.2 with 27% decayed teeth in 1963-70 and 3.1 with 17% decayed teeth in 1988-94 (White, 1995).

The most recent national examination survey (NHANES III, which was conducted from 1988 through 1994) found significant differences in children's DFMT scores with higher scores among older children, ethnic and racial minorities, and low-income children (Vargas, 1998). This finding also held for the scores for dental surfaces among primary and permanent teeth, the measure used in the DRM as an indictor of restorative dental care needs. Appendix 1 lists the decay levels by the eighty population subgroups for the baseline decay rates and the expected new decay rates. The NHANES III uses the Mexican American group as the only identified Hispanic population subgroup, with other Hispanics populations placed in the "Other" category. Since many states do not have a count of this subgroup, the users will have to determine the most appropriate way to input their Hispanic population.

The DRM – Design, Inputs, and Outputs

The DRM is a spreadsheet model that estimates the dentist requirements (general and pediatric dentists) for dental care using a backlog and maintenance component of children's care. For each of eighty population subgroups, the dental care need is based on one check-up per year and the estimated rate of decayed tooth surfaces (primary and permanent) that need filling, with rates that differ for each of eighty population subgroups. The NHANES III data is used to estimate the decayed surfaces at baseline and new decay for children from the eighty subgroups derived from five age groups, four ethnic/racial groups, and four family income levels.

The model allows users to set a target for the percentage of decayed surfaces that will be filled (default value of 84% of tooth surfaces filled) a rate from NHANES III for population group with the best treatment scores. The annual check-up rate has a default value of 90% of children.

The dentists' productivity is expressed as the time needed per service expressed as parts of an FTE based on 2,000 hours/year. This is estimated for each of three services – initial check-up, follow-up check-up, and filling a decayed surface. Default values are 30 minutes for initial check-up, five minutes for follow-up check-up, and fifteen minutes for filling a dental surface. The model uses an estimate of existing dentists' excess capacity to provide care (default value of 1%), under an assumption that there is a pool of dentists that have unused capacity which when pooled together is the equivalent of 1% of the FTE dentists. The model allows for the dental care workload to be split between pediatric dentists and general dentists, with default values set at the following for pediatric dentists - 100% of care for children under three years of age and care for 6.6% of the children for each of the remaining age groups.

The user enters the following.

- the number of children in each of eighty population subgroups based on
  - five age groups: 0-3 years, > 3 to 6 years, > 6 to 10 years, > 10 to 13 years, and >13 to 18 years,
  - four ethnic/racial groups: Mexican American, Non Hispanic (NH) African American, NH White, and Other;
  - four family income groups based on the federal poverty levels (FPL): 0 – 99% FPL, 100-149% FPL, 150-199% FPL, and 200+% FPL.
- four estimates for dentists
  - the current supply of both pediatric and general dentists,
  - the excess capacity estimates for each dentist group,
  - the workload met by each dentist group (general dentists and pediatric dentists),
  - the FTE time allotted for each of three procedures (initial and follow-u[p check-up and filling a decayed surface)
- target values for
  - percent of decayed surfaces that will be filled, and
  - percent of children who will receive a check-up.

The output is presented as an estimate of the requirements for dentists (FTE general dentists and pediatric dentists) to provide the backlog and maintenance care. The dental FTE is apportioned between existing dentists and new dentists. The current model does not allow any adjustment for the estimated percent of dentists who will provide care to children with Medicaid or CHIP coverage, although this is a planned enhancement.

State-Level Application of DRM – Illinois

The DRM will be applied to Illinois data, with the following information provided as background to be used to assess the model output. The supply of Illinois dentists is shown in Table 2, based on data obtained from the ADA and Illinois Medicaid program. Note that the supply of active patient care general and pediatric dentists is 6,061 with only 140 pediatric dentists (about 2.3% of dentists). Data from the Illinois Medicaid program (which in Illinois includes the CHIP enrolled children) shows that only 2,037 of these dentists have signed up to be Medicaid/CHIP providers, with 1,594 having submitted at least one claim, and only 740 having submitted more than 100 claims. Thus only about 26% of dentists provided any care and only 12% provided the equivalent of care to more than 2 children per week.

Table 2 Illinois Dentist Information, 1999/2000

Active Patient Care Dentists (General and Pediatric dentistry)	6,061
General Dentists	5,921
Pediatric dentists	140
Medicaid/CHIP Participation by Dentists	
Enrolled as provider	2,037
Submitted one or more claims/year	1,594
Submitted one hundred or more claims/yr	740

There were just over 1.0 million Illinois children with family income levels under 185% of poverty, the upper income threshold for CHIP eligibility. (The population of the State is approximately 12.8 million). The estimated 1.0 million children includes all children in this income category, regardless of insurance coverage (e.g. eligible and enrolled in Medicaid or CHIP and privately insured children). The Illinois children were apportioned into the two family income levels of the model, (0-99% and 150-199%), most closely aligned to the available data for Illinois. The ethnic/racial breakdown of the NHANES data was applied to the actual counts of Illinois children by age, since detailed information is not currently available for Illinois children. (See the DRM model summary, in appendix 2, for this information).

The model was run using all default values. The printout of the model inputs and output is shown in appendix 2. The model estimates of the number of dentists to provide the backlog care at entrance into a program for the 1.0 million children to be the equivalent of 384 FTE dentists (split as shown between general dentists and pediatric dentists). The model estimates that existing dentists could provide the equivalent of 60 FTE dentists (entered by user as excess capacity o existing dentists) and that the net new requirements would be 324 FTE dentists. Realistically, it is expected that the backlog dental care would be spread over several years as children entered the program in an incremental fashion.

For maintenance care of children in the program, the model estimates that 103 FTE dentists would be required. It estimates that the equivalent of 60 FTE dentists could be obtained from existing dentists and that 43 new dentists would have to be added to the state dental workforce.

The following discussion will focus on the maintenance requirements. For the discussion, we will not assume an excess capacity, so all needed dentists will be new. Several considerations of the model design and assumptions will be discussed to assess the model outputs. First, the 103 FTE dentists per 1.0 million children would yield about 1 dentist per 10,000 children – a high number of children per dentist. As a point of reference, an area may be designated as a dental shortage area if the population to dentist ratio is higher than 4,500 population to one dentist.

The model estimate for dentist requirements reflects the relatively low intensity of care that the model assumes for maintenance care – one check-up visit (with an estimate of five minutes of dentist time) and an average of less than 1.0 dental caries surface to be filled per child (with a time estimate of 15 minutes of dentist time). The model also assumes a 2,000 hour work year for dentists, a high estimate based on ADA surveys. In addition, the time estimates for the dental services may not be sufficient for Medicaid and CHIP children, where their high needs may require greater time estimates.

This model does present an estimate for state-level planners that can be modified with changes in existing user inputs to the model. The estimate of about 100 FTE dentists may be an underestimate, but is expected to be in a range considered as a reasonable estimate.

Model Enhancements
The model developers (VRI) and HRSA are planning to revise the model and to add model enhancements. These are expected to include some of the following. With regard to the population groups, the family income groups may include a category at 133% of FPL since this is a common cutoff for Medicaid programs, and consequently the CHIP eligible groups.

With regard to dental productivity several changes are being considered. The estimates of work hours per FTE dentist will be revised to the ADA survey findings. Time estimates for services are currently user inputs, and users may be encouraged to consider modifying their inputs based on staffing information (e.g. dental hygienists in dental practices) or increasing the time allotted for special care needs of the children, or feedback from practicing dentists on their time allotments.

The backlog concept will be revised to recognize that changes in children's utilization will be incremental. Many states have set targets for changing the utilization rates for their children over three to five years. The ability of states to recruit dentists to provide care for the CHIP and Medicaid children will require both recruitment of new dentists and recruiting greater numbers of existing dentists to participate in these programs. The dentists' participation will be affected by the states' reimbursement rates. The model may consider a way to allow users to include reimbursement rates, commonly expressed as the percent of the usual, customary, and reasonable rate, or a percent of the rates established by dental reference groups.

Conclusions
The DRM represents a potentially useful tool for state level planning for increasing access to dental care and estimating the dental workforce needs. It comes at a critical time as many states have recognized their poor performance with low dental utilization rates among children with Medicaid. States are taking steps to try and improve access through greater program participation by dentists. The model will allow for estimates of the requirements of dentists with many variations in assumptions and inputs that can be tailored to the circumstances within the state. Further model enhancements will improve the usefulness of the model.

Acknowledgements
This study was funded by the Bureau of Health Professions, Health Resources and Services Administration, project officer Louis Kuta.

References

American Dental Association. *ADA Dental Workforce Model: 1997-2020.* Chicago. 1999.
American Dental Association. *The 1998 Survey of Dental Practice.* Chicago. March 2000.
Brown, L.J.; Wall, T.; Lazar, V., Trends in Untreated Caries in Permanent Teeth of Children 6 to 18 Years Old, JADA, Vol. 130, November 1999
Brown, LJ, Wall, T, Lazar, V, Trends in Total Caries Experience: Permanent and Primary Teeth, *JADA*, Vol. 131, February 2000
Byck GR. A Comparison of the Socioeconomic and Health Status Characteristics of Uninsured, SCHIP-Eligible Children in the United States to Other Groups of Insured Children: Implications for Policy. *Pediatrics* 2000;106:14-21.
Center for Research and Public Policy. *Connecticut Medicaid Study.* New Haven, CT: Center for Research and Public Policy. 1999.
Edelstein BL, Douglass CW. Dispelling the myth that 50 percent of U.S. schoolchildren have never had a cavity. *Public Health Reports.* 1995; 110:522-530.
Edelstein B, Manski R, Moeller J. Pediatric dental visits during 1996: An Analysis of the federal medical Expenditure Panel Survey. *Pediatric Dentistry.* 2000; 22:17-20.
Nainar SM, Edelstein B, Tinanoff N. Access to dental care for Medicaid children in Connecticut. *Pediatr.Dent.* 1996; 18:152-153.
Smith, VK. *CHIP Program Enrollment: December 1998 to December 1999.* The Kaiser Commission on the Uninsured. July 2000
United States General Accounting Office. *Oral Health: Dental Disease is a Chronic Problem Among Low-Income Populations.* GAO/HEHS-00-72, April 2000
U.S. Department of Health and Human Services. *Oral Health in America: A Report of the Surgeon General.* Rockville, MD: U.S. Department of Health and Human Services, National Institute of Dental and Craniofacial Research, National Institutes of Health, 2000.
U.S. Department of Health and Human Services, Health Resources & Service Administration. *United States Health Workforce Personnel Factbook.* 1999.
Vargas, C, Crall, J; Schneider, D., Sociodemographic Distribution of Pediatric Dental Caries: NHANES III, 1988-1994, *JADA*, Vol. 129, September 1998
Venezie RD, Vann WFJ. Pediatric dentists' participation in the North Carolina Medicaid program. *Pediatr.Dent.* 1993; 15:175-181.
White BA, Caplan DJ, Weintraub JA. A quarter century of changes in oral health in the United States. 1995 *Jour Dent Ed*, 59:19-57.

Appendix 1. Decay levels by Population Category for the DRM

	Decayed Surfaces	Filled Surfaces	New Decayed Surfaces
100-149%			
0-3 years	2.62	0.28	0.02
>3-6 years	3.71	2.23	1.12
>6-10 years	2.96	2.57	0.39
>10-13years	0.7	1.83	0.26
>13-18 years+A10	1.33	4.17	1.16
150-199%			
0-3	1.28	0.1	0.16
>3-6	3.56	4.01	0.88
>6-10	0.88	3.08	0.33
>10-13	1.07	1.71	0.47
>13-18	0.93	3.83	0.82
	NH African-American		
100-149%			
0-3	1.15	0.22	0.02
>3-6	1.74	2.05	1.12
>6-10	1.94	1.85	0.39
>10-13	1.99	1.47	0.26
>13-18	1.65	3.79	1.16
150-199%			
0-3	1.12	0.05	0.16
>3-6	0.96	1.38	0.88
>6-10	1.86	0.35	0.33
>10-13	0.88	0.71	0.47
>13-18	1.48	2.81	0.82
	NH White		
100-149%			
0-3	1.35	0	0.02
>3-6	2.33	2.81	1.12
>6-10	1.04	4.26	0.39
>10-13	0.47	2.18	0.26
>13-18	0.83	4.68	1.16
150-199%			
0-3	0.83	0	0.16
>3-6	0.92	1.19	0.88
>6-10	0.93	3.68	0.33
>10-13	0.96	1.58	0.47
>13-18	1.47	3.52	0.82
	Other		
100-149%			
0-3	4.05	0	0.02
>3-6	3.62	2.12	1.12
>6-10	0.58	5.2	0.39
>10-13	0	3.62	0.26
>13-18	6.92	5.4	1.16
150-199%			
0-3	0	0	0.16
>3-6	7.71	4.99	0.88
>6-10	0.96	7.53	0.33
>10-13	0.12	3.7	0.47
>13-18	0.06	5.87	0.82

INSTRUCTIONS

1. In the POPULATION VALUES table, enter the population data specific to age, Federal Poverty Level (FPL), and Race.
2. Modify the INPUT PARAMETERS data for your specific circumstances.
3. View results in the RESULTS table.
4. Tabs are named as: Federal Poverty Level_Race/Ethnic Group
5. NH = "Non-Hispanic"
6. White cells indicate input data. Red text indicates totals.

AGE GROUPS

1 : 0 - 3 years
2 : >3 - 6 years
3 : >6 - 10 years
4 : >10 - 13 years
5 : >13 - 18 years

POPULATION VALUES

FPL	Age Group	Mexican American	NH African American	NH White	Other	Total
0-99%	1	19,883	61,565	24,960	2,140	
	2	29,061	72,271	44,413	3,913	
	3	33,653	93,559	68,067	2,864	
	4	24,829	74,441	34,129	2,061	
	5	33,143	96,441	54,901	4,641	
100-149%	1					
	2					0
	3					0
	4					0
	5					0
150-199%	1	11,883	5,742	18,007	722	36,354
	2	15,261	12,526	19,029	2,305	49,121
	3	17,656	13,077	27,592	3,143	61,468
	4	11,437	11,156	21,539	2,737	46,869
	5	26,407	18,105	32,289	4,865	81,666
200+%	1	0	0	0	0	0
	2	0	0	0	0	0
	3	0	0	0	0	0
	4	0	0	0	0	0
	5	0	0	0	0	0
Totals		223,213	458,883	344,926	29,391	1,056,413

INPUT PARAMETERS

Target Percentage for Filled Surfaces	84.00%
FTEs per Initial Checkup	0.0002500
FTEs per Follow-up Checkup	0.0000416
FTEs per Filling	0.0001250
Checkup Target Value (%)	90.00%
Current Supply of Dentists	
for General Dentists	5,921
for Pediatric Dentists	140
Excess Capacity (%)	
for General Dentists	1.00%
for Pediatric Dentists	1.00%
Workload Met by Pediatric Dentists (%)	
for Age Group 1	100.00%
for Age Group 2	6.60%
for Age Group 3	6.60%
for Age Group 4	6.60%
for Age Group 5	6.60%

RESULTS

	Backlog		Maintenance	
	Required General Dentist FTEs	Required Pediatric Dentist FTEs	Required General Dentist FTEs	Required Pediatric Dentist FTEs
Total New Requirements	308	76	89	14
Existing Extra Capacity	59	1	59	1
Net New Requirements	249	75	30	13

FORECASTING THE PHYSICIAN WORKFORCE

Richard A. Cooper, M.D.
Health Policy Institute
Medical College of Wisconsin
Milwaukee, Wisconsin 53226
rcooper@mcw.edu

I. QUANTITATIVE MODELS

In 1933, the Committee on the Costs of Medical Care (CCMC) published its historic treatise entitled, *The Fundamentals of Good Medical Care*, describing the dimensions of the physician workforce in precise, quantitative terms (1). The CCMC's approach systematically measured the prevalence of disease, determined the exact number of physician encounters required for each and designated the time (in minutes) for each encounter. Its unique and enduring contribution was to establish two basic tools for workforce analysis that dominated thinking for the remainder of the 20th Century: reconstructing the system from its component parts and quantitating the parts using the *metric of time.*

Applying these tools, the CCMC concluded that, in the aggregate, good medical care in 1929 required exactly 283,131 hours of physician time, which they equated to 140.5 physicians per 100,000 of population, a figure that was 10% greater than the existing supply.

Almost half a century later, the Graduate Medical Education National Advisory Committee (GMENAC) reached into the past for a model that it could use to determine the number of physicians that were required in each of the specialties (2). While retaining the CCMC's core methodologic tools, it modified the approach to create its "*adjusted needs model*." However, like the earlier model, its dependence on disaggregating and reconstituting the universe of care, coupled with its need to assign the metric of time to both the elements of care and the effort of physicians, seriously handicapped its ability to determine what actually was occurring. But GMENAC went one step further. It proceeded to extrapolate its calculations twenty years into the future, predicting that there would be a surplus of 145,000 physicians (30%) in the year 2000. Although this prediction proved to be excessive, it has had a pervasive and continuing influence on health policy discussions.

With the increasing availability of data about clinical practice in the early 1990s, GMENAC's successor, the Council on Graduate Medical Education (COGME), adopted the *demand-utilization model* for workforce planning (3). Rather than relying on epidemiologic data, it assessed the need for physicians based on actual measurements of services provided, drawing upon the resources of national databases such as the National Ambulatory Medical Care Survey and Medicare claims data. However, like its predecessors, it attempted to recreate physicians from their component tasks and to standardize them by applying the metric of time, and it, too, failed. For example, only six years ago, the COGME projected that there would be a surplus of 80,000 physicians in the year 2000, including a 47% surplus of specialists (4).

As managed care emerged, a new avenue of analysis, the *requirements model*, appeared. It was based on physician utilization in staff/group model HMOs. These seemingly "closed systems" should, it was reasoned, be able to account for all of the care provided and all of the time necessary for physicians to provide it. However, the HMOs from which this model was built represent a small and shrinking segment of clinical practice, and the assumptions and extrapolations required to describe the entire system from this narrow pedestal are complicated and tenuous. As a result, the conclusions have been far from the mark. Indeed, in what was characterized as "the most complete forecast to date," carried out on behalf of COGME in 1994, Weiner predicted a surplus of 165,000 physicians (30%) in the year 2000, including a 64% surplus of specialists (5). Combined with COGME's earlier predictions, these projections led to a call for the closure of 20 US medical schools, a sharp decrease in specialty training and the curtailment of funding for international medical graduates, measures that were partially addressed in the Balanced Budget Act of 1997.

Thus, beginning with the CCMC's report in 1933 and continuing through GMENAC's in 1980 to COGME's various reports and studies in the 1990s, assessments of the physician workforce have been dominated by a linear, mathematical mode of thinking based on *dissecting and reconstituting the health care* system and standardizing its components according to the *metric of time*. The errors associated with applying this process to a multiplicity of diseases, an array of services and a diversity of both patients and physicians are enormous. Using it to project future needs further compounds the error, often in ways that are not apparent in the final product. Indeed, it seems clear that physician surpluses in the range of 15-30% that were projected by these quantitative methods for the year 2000 are not *consistent with the current realities.*

II. THE "TREND MODEL"

The Trend Model, presented below, offers an alternative to the "quantitative models" discussed above. It is constructed around the principles

		Adequacy ↱ of services		↱ Economy ↱ Culture ↱ Population		↱ Fiscal		↗ **DEMAND**
SUPPLY	x	**SUFFICIENCY**	x	**TRENDS**	x	**CONSTRAINTS**	⇒	**FUTURE**
		↳ Utilization of providers		↳ Productivity ↳ Substitution ↳ Attrition		↳ Training		↘ **SUPPLY**

of assessing the trends that affect the supply of physicians and the demand for their services.

The dominant trend is the economy. Even in 1933, the CCMC recognized that "compelling economic forces" influence the number of *physicians (1). These forces act not only in a* direct way but also indirectly by influencing the development and utilization of technology and the structure of systems of health care delivery and financing.

The other major trends influencing demand are population growth and cultural attitudes toward health care. Trends that influence supply include physician productivity and attrition and the provision of "physician services" by nonphysician clinicians (NPCs). In addition, the imposition of external constraints, through *controls on training or financing, may, at least in* the short term, override the natural evolutionary processes.

SUPPLY

The starting point in the Trend Model is an estimate of the physician labor force. The year 1990 has been taken as the "base year" for this and other elements of the model. Therefore, physician supply is estimated from 1990 forward. All active physicians are counted, irrespective of the nature of their activity or the extent of their work effort. This recognizes that physicians serve varying roles and that the mix of roles and time commitment to each change over time. These roles include not only the traditional ones of direct patient care, teaching, research and administration but other roles such as participation in pharmaceutical, biotech and medical equipment companies; medical direction of insurance companies, health plans and managed care organizations; roles in professional organizations, regulatory agencies and public health departments; and others. It also recognizes that physicians differ in the time that they devote to professional activities and in the efficiency with which they accomplish their professional tasks. Therefore, physician supply is expressed as a head count rather than as a derived number of FTE physicians related to *certain tasks. Measures of physician supply are* obtained from sources such as the AMA Master File, specialty society records, recertification data, etc. Differences among the data from these various sources (which are common) are reconciled in order to make final estimates.

SUFFICIENCY

The level of supply that is estimated in this manner cannot be taken as a normative value from which future supply is projected. Rather, this level must be interpreted in the context of the *utilization of physicians* (job opportunities, desire for additional workload, etc.) and the *adequacy of services* being provided (waiting times, unmet needs, excessive services, etc.). Information regarding physicians is derived from surveys and consensus panels and from data provided by group practices and other organizations that employ physicians. Several professional societies routinely conduct such surveys of their members, and some also survey graduating residents. Information concerning patients' perceptions of the adequacy of physician supply are obtained from the National Health Interview Survey and surveys performed by public policy and consumer opinion organizations.

TRENDS

Projections of the future physician labor force are based on six major trends. Three (productivity, attrition and substitution) directly affect the available supply of services. Two (economy and culture) are the pillars of future demand. The final one (population) is both fundamental to demand and intrinsic to the model, which expresses both supply and demand in per capita terms. However, before discussing these six trends, two trends that are commonly associated with the demand for physicians but that are not separately included in the Trend Model require comment. These are *technology,* and the *aging population.*

Technology is not separately considered because it is principally a function of the *economy*. An expanding economy has the resources to invest in technology, and a prosperous nation has the resources to purchase the products of technological development. However, while associated most strongly with economic trends, the growth of technology influences other trends. For example, some technologies facilitate the *substitution* of generalists for specialists or of NPCs for physicians. In addition, the prominence of technology, coupled with the promise of future technologies, contributes to a *culture* that is willing to devote increased resources to health care. Thus, although not separately considered, technology in prominent in the Trend Model.

Aging of the population creates a reservoir of disease and disability that demands medical care. In some cases, this represents a net increase in *demand,* while in others it is the deferment of care that otherwise would have been provided at a younger age. But, as with technology, the quantity of care that results is ultimately determined by the resources that are available (6), which, in turn, depend on the state of the economy. Therefore, aging is not separately considered.

Population Trends

Population is a critical component of the Trend Model. Data and projections regarding population are derived principally from the Bureau of the Census. Unfortunately, political considerations require the Bureau to under-report the US population. Moreover, the trend has been for the Bureau to increase its projections of the future population over time.

Population trends depend primarily on birth rate and immigration. There is a great deal of uncertainty regarding future birth rates, particularly since they differ among ethnic groups. For example, the birth rate of the Hispanic population, the most rapidly growing segment of the US population, has tended to be higher than the norm, but it is uncertain whether this will continue or whether Hispanics (and certain other immigrant groups) will adopt the lower birth rates of the population overall.

Immigration has been constrained in recent years. However, the current labor shortage, coupled with a falling ratio of workers to retirees at a time when there are growing populations in many less developed countries, is leading to calls for more immigration. Therefore, the population estimates applied to the Trend Model have been modified upward from those of the Census Bureau to adjust for under-reporting and to include the likelihood of higher rates of immigration over the coming years.

Productivity Trends

Productivity is influenced by both the professional time and work output of physicians. Among the trends influencing productivity are gender, age, life-style, employment status and efficiency. These are not independent variables but, rather, are interconnected. The Trend Model assesses the impact of these various trends on overall productivity relative to the productivity of physicians in 1990. In addition, because residents account for such a large portion of the physician workforce, adjustments are made for their productivity relative to that of fully trained physicians.

Gender: Women physicians have tended to work approximately 15% fewer hours and to see 15% fewer patient visits than male physicians. It is assumed that the same differences apply to the nonclinical roles that physicians serve. More than 40% of current residents are women, and the Trend Model assumes gradual transition to a physician workforce that is almost 50% female.

Physician age: On average, physicians over the age of 55 work 10% fewer hours than physicians who are less than 45 years old. The average age of physicians is increasing as the cohort of young physicians that was generated by medical school expansion in the 1970 comes into equilibrium. Beginning in 2010, the number of new physicians will approximately equal the number leaving the workforce due to death and retirement.

Life-style: There is a trend for all physicians, male and female, to work fewer hours. This is attributed to their greater emphasis on personal time. It is assumed that this trend will continue.

Employment: Physicians who are employees of organizations tend to work fewer hours than physicians who are self-employed or who have an ownership position in their organization. The trend has been for an increasing percentage of physicians to be employed. This may relate in part to life-style and professional considerations, but it is also influenced by the need to capitalize clinical practices.

Efficiency: There is a broad trend toward increased productivity in the US labor force, related principally to information and technology. It appears that medical care has not shared in this increase to the extent experienced in other sectors of the economy. However, it is likely that it will as the use of computerized medical records and other tools of information management become more prevalent and as the technology of monitoring and communicating with patients advances. To some extent, this trend counterbalances those described above.

Residents: Residents account for approximately 15% of active physicians. However, their work effort is less. Previous studies have assumed that the productivity of residents is 35%-75% that of a fully trained physician. The Trend Model counts resident effort in various specialties at 40-70% of the effort of practicing patient care physicians (7).

Attrition Trends

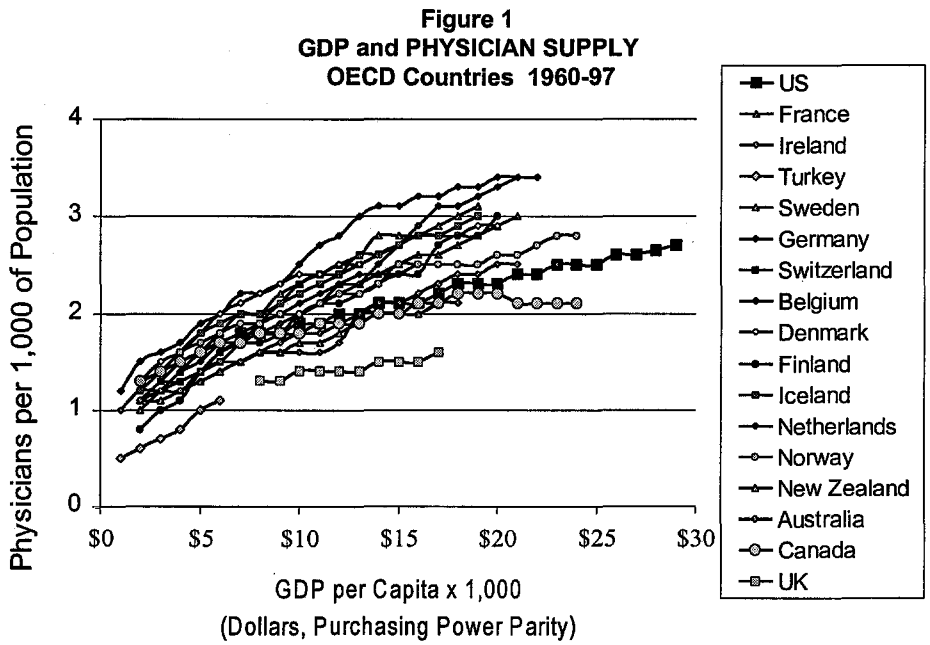
The Trend Model includes separate trends for death rates and retirement. Death rates are taken from actuarial tables. The major variable is attrition. Trends in attrition are assessed through information obtained from surveys (as periodically conducted by the AMA, recruiting firms and professional associations), from recertification data and from the membership records of professional societies. Recent surveys indicate that physicians are leaving their professional roles at earlier ages and that they are more likely to do so in the future. Trends in the attrition of NPCs are assessed using data obtained from surveys conducted by the BHPr and by the relevant professional associations.

Substitution Trends

"Nonphysician clinicians" is a term applied to a group of licensed professionals who have in common the authority to be the point of first contact for patients, to take the principal responsibility for the care of patients (under at least some circumstances) and to provide elements of care that fall within the spectrum of "the practice of medicine." These professions include nurse practitioners (NPs), clinical nurse specialists (CNSs), certified nurse-midwives (CNMs), physician assistants (PAs), nurse anesthetists, optometrists, podiatrists, psychologists and the alternative disciplines of chiropractic, acupuncture and naturopathy.

A confluence of dynamics has propelled the growth of many of these disciplines, both in numbers of practitioners and in their licensed scope of practice (8, 9). At the same time, technology has allowed previously complex procedures to become safer and more readily delegated to NPCs, and system changes have further facilitated the distribution of responsibility from physicians to NPCs. The growth limits of this phenomenon are not clearly defined, but the trends seem clearly established.

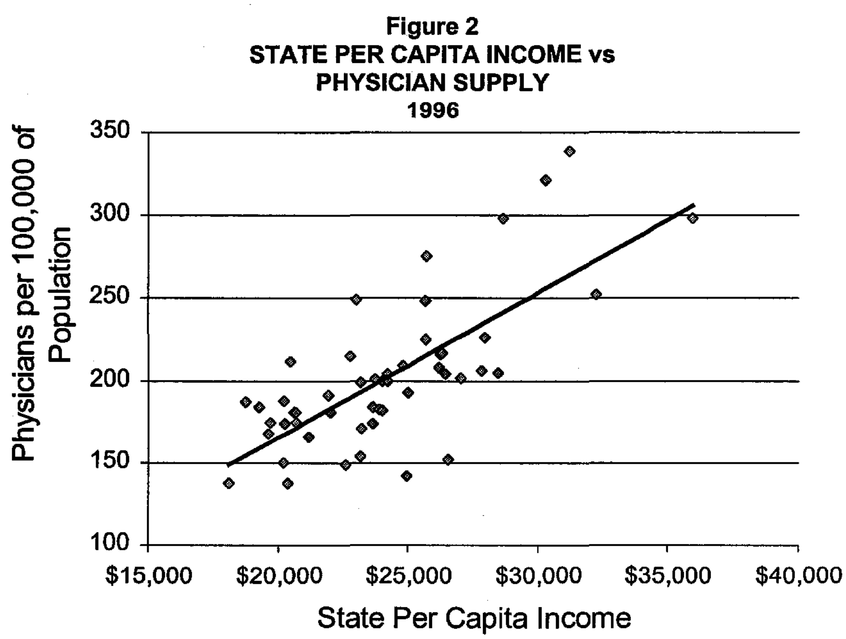
While there is increasing overlap between physicians and NPCs, the work-scope of NPCs does not fully overlap that of physicians, nor do NPCs collectively encompass the range of practice of physicians. Rather, they tend to treat conditions that are less complex and to provide services that are more routine. Moreover, NPCs generally work fewer hours than physicians. Therefore, the substitution of NPCs for physicians is not on the basis of a simple head count. Rather, specific *substitution ratios* take into account the degree of overlap, the comparative hours worked and the efficiency of delivering services. These ratios are becoming larger as the training and licensed authority of NPCs expands (9).



Economic Trends

The dominant factor in the growth of demand for physicians is the overall growth of the economy, as measured by indices such as the gross domestic product (GDP), personal consumption and disposable income. Figure 1, which is derived fro the Organization for Economic Cooperation and Development (OECD), demonstrates this trend among seventeen member nations over the period from 1960 to 1997. Excluded from this analysis are Japan, with a system that bears little resemblance to that of the other countries, and the four Mediterranean nations (Italy, Greece, Spain and Portugal) that produce physicians well beyond their capacity to utilize them.

The relationship between physician supply per capita and GDP per capita is similar among these countries. Even Turkey, whose per capita GDP in 1997 was less than that of the US in 1960, follows the same trend line. However, there are two important exceptions. The first is the UK, which has traditionally constrained



physician supply and now faces a physician shortage. The second is Canada, which began to constrain physician supply in the early 1990s and is also experiencing a physician shortage.

While these trends have similar slopes they display different absolute magnitudes of supply at any level of GDP. This may relate to differences in work effort of physicians among countries. However it also may relate to differences in culture and mores.

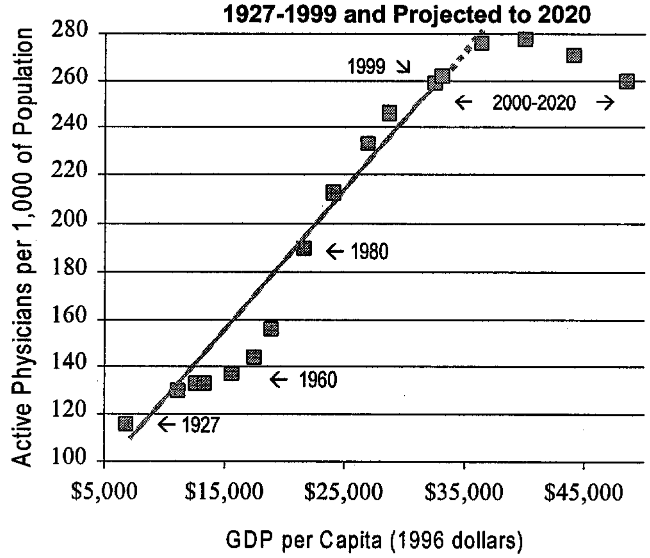
A similar trend was observed when the per capita income of all 50 states was compared with the per capita supply of physicians for a single year (1996) (Figure 2). Moreover, when the gross state product (GSP) of each state was plotted against the state's per capita physician supply over the period from 1983 to 1997, a family of trend lines similar to those depicted in Figure 1 was obtained (data not shown). Like the international comparisons, there were important exceptions. California, Arizona and Nevada followed a pattern similar to that of Canada and the UK, with relatively flat trend lines over the fourteen-year period of observation.

Figure 3 shows a more detailed representation of the relationship between GDP and physician supply in the US over the period of 72 years from 1927 to 1999. This analysis utilized economic data from the Bureau of Economic Analysis (BEA) and data on the supply of active physicians from the Bureau of Health Professions (BHPr). Also shown is the projected supply of physicians during the period from 2000 to 2020, as published previously (10). This is plotted against a projected per capita GDP that follows an annual growth trend of 2.0% per in inflation-adjusted dollars.

A number of observations can be made from Figure 3 that are relevant to the Trend Model. First, a general relationship between GDP and physician supply can be traced back to 1927. Second, during the period between 1940 and 1965 there were fewer physicians per capita than would have been predicted. This coincided with a growing perception of a physician shortage that culminated in federal legislation, leading to an expansion of US medical schools and a relaxation in the immigration barriers for foreign physicians. Physician supply was re-established at the trend line by 1980 but deviated in the direction of oversupply in the early 1990s before returning to the trend line in 1999. This is consistent with the current perception that, despite pockets of over-supply and under-supply, physician supply and demand are in balance (7, 11). Finally, the period from 2000 to 2020 recapitulates the earlier period of a physician shortage that was experienced between 1940 and 1960.

From these and other analyses, a relationship between GDP and physician supply was defined that predicts that for every 1.0% increase in GDP per capita there will be a 0.6% increase in physician supply per capita. This is less than the national income elasticity of health

Figure 3
GDP per CAPITA vs.
ACTIVE PHYSICIAN SUPPLY
1927-1999 and Projected to 2020



care, which is approximately 1.5% (12, 13).

While the relationship between GDP and physician supply pertains to physician supply overall, it does not apply equally to the various specialties. For example, general/family practice displays no such relationship. Indeed, the ratio of primary care physicians to population has been constant for 50 years. The steepest slope is displayed by the medical subspecialties, while surgery is intermediate.

The relationship between GDP and physician supply that was developed from Figure 3 and related studies is projected as the "GDP Demand Trend" in Figure 4. This represents the demand for "physician services" irrespective of whether these services are provided by physicians or NPCs.

Cultural Trends

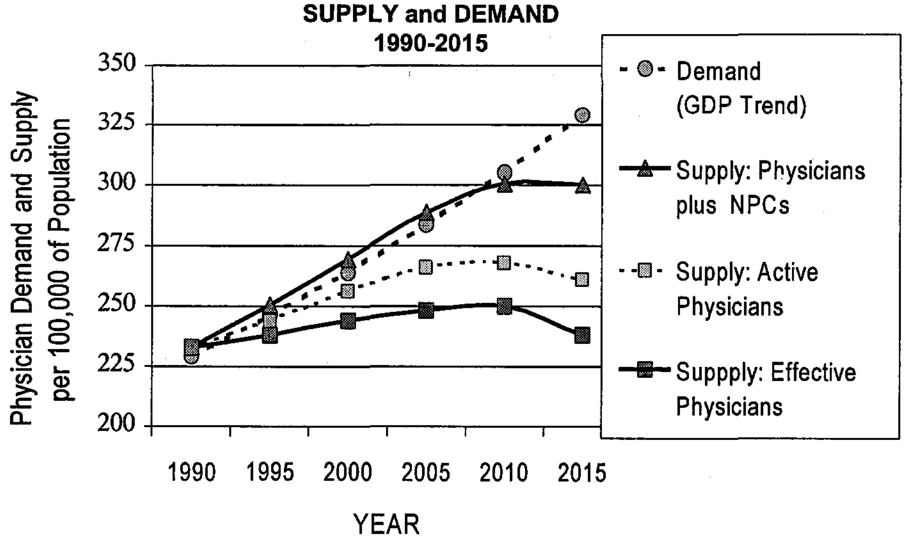
The international data in Figure 1, as well as similarly constructed state data described above but not shown, display parallel trend lines relating physician supply to GDP (or GSP), each with a high correlation coefficient over long periods of time. However, the absolute level of supply in the various countries or states differs at each level of GDP (or GSP). As noted above, this may be due to differences in the work effort of physicians, particularly among countries. However, it also may relate to differences in cultural values and expectations and in the way that communities organize their health services.

The level of health care expenditures and of physician supply in each geopolitical region appears to be determined by the blending of its economic potential with the "vision of a good society" held by its citizens (14). This blending engages the natural tension between public policy, capital markets, governmental regulation and individual action, a process that Arrow has termed the "social adjustment toward optimality" (15). It is this process that ultimately governs resource allocation and income redistribution. The striking observation with respect to the relationship between GDP and physician supply is how stable these relationships are within each region over long periods of time.

CONSTRAINTS

In using trends to project the future, it is assumed that there will be a natural evolution of the current fiscal and organizational characteristics of the health care system and of the societal fabric in which it exists. These characteristics include an emphasis on technology and specialization, a responsiveness to consumer demand and an expanding portion of the GDP devoted to health care. While some have championed all of these as desirable, others have urged a reversal of the current trends by slowing technology, increasing the emphasis on primary care, curtailing consumer demand and

Figure 4
PHYSICIAN and NONPHYSICIAN CLINICIAN SUPPLY and DEMAND 1990-2015



redirecting national spending to other priorities.

Attempts have been made to control costs, either by limiting the volume of service or the *level of payment per unit of service. Supply* constraints have been introduced through measures to restrict medical education, such as those in the Balanced Budget Act of 1997 and in the Canadian measures to decrease class size earlier in the 1990s. The results of such constraints are apparent in the deviation from the norm of the UK, Canada and California. While the Trend Model is constructed around observed trends, it also permits the introduction of fiscal and supply constraints analogous to those mentioned.

Although there are many examples of constraints on health care spending and on the training of physicians, the time-frame of these constraints has tended to be short, rarely encompassing as much as 10-20 years. Moreover, as evident from Figure 3, constraint tends to be followed by excess, as the actual supply of physicians moves around the trend line over long periods of time. Ultimately, the supply demand equilibrium is re-established at levels that appear to correspond to predictions based on economics, culture and demographics. Therefore, the use of constraints in this model is most applicable to short-term projections.

FUTURE

The Trend Model leads to a calculation of future physician *supply* and the *demand* for physician services that are a consequence of the various trends that are considered above.

Supply

Future physician supply is expressed as the number of active physicians who will be in the labor force relative to the base year of 1990. For purposes of the model, it is assumed that 22,000 new physicians will be trained annually, as has been the case over the past decade. The future supply of physicians is extrapolated based on the number who are now active, the number newly trained and the number who will leave the profession due to death and retirement. This number is further modified by the trends in physician productivity and in the supply and substitution of NPCs, as described above. These various adjustments lead to a calculation of the magnitude of the effective labor force (including both physicians and NPCs) relative to the magnitude of this labor force in 1990.

Demand

The term "demand" is used to describe the projected size of labor force that will be required in order to deliver the quantity of service that is predicted, based on the economic, cultural and population trends described above. As is true for supply, future demand is expressed relative to the number of active physicians per capita in 1990. It is this derived number that forms the basis for decisions concerning changes in the numbers of students and residents who must be trained in order to create a supply that satisfies this future demand.

Limitations of the "Trend Model

Like the "quantitative models" described earlier, the Trend Model applies a common metric. However, rather than applying the mathematical *metric of time* to diseases, visits and providers, it depends on an analysis of the *trends* that affect the provision and utilization of medical services. As a result, the various assumptions used are not immersed within a multiplicity of time assignments but, rather, are open and accessible, thereby facilitating their modification or reinterpretation. The error of this model is fundamentally a product of the errors of the individual trends, and these errors become magnified as the time projected lengthens. Moreover, in applying this model, the time frame of the trends considered must be long in relation to the time-frame of the extrapolations being made (13). Near-term projections (3-5 years) can depend on short-term trends, but projections that are within the time-frame of importance to training decisions (10-20 years) require trends that span many years.

III. APPLICATION OF THE TREND MODEL

Figure 4 displays an application of the Trend Model to an analysis of the physician workforce over the past decade and a projection to the year 2015. This is a multi-step process.

Active physician supply: The first step is a representation of the projected supply of active physicians. The curve shown in Figure 4 was constructed based on a constant input new physicians, a discounted effort by resident physicians and trends in attrition and population, as described above.

Effective physician supply: The second step is the translation of active physician supply to "effective physician supply" by applying the various trends in productivity discussed above and published previously (10). The actual calculation applies the decremental effort since the base year 1990.

Effective supply of physicians and nonphysician clinicians: The effective physician supply derived in step two is modified by the additional contribution made by NPCs. In a manner similar to the calculation of the decrement productivity, this calculation of NPC effort represents the incremental effort since the base year 1990. The contribution of each NPC discipline is based on the projected number of practitioners (8) and substitution ratios for each. These range from 0.1 for optometrists to 0.7 for nurse anesthetists. For most discipline, substitution ratios are increasing over the period projected, based on trends in their practice prerogatives (9). This combined supply of effective physicians and NPCs represents the projected labor force devoted to "the practice of medicine," as practiced by physicians.

GDP Demand: The demand for physician services in the future is projected based on the assumption that there will be a continuation of the trends that relate GDP to health expenditures (12, 13, 16) and to physician supply (Figs. 1-3). The relationship that was applied to the model (0.6% increase in physicians per capita for each 1.0% increase in GDP per capita) was derived from Figure 3. It is further supported by a larger body of data on economic correlates at the state, national and international levels, each spanning 15-35 years.

Supply-demand relationships: The data and projections presented in Figure 4 indicate that, in absolute terms, there has been a shortage of physicians since the early 1990s. However, manifestations of this shortage were averted by the training and licensure of a growing number of NPCs. In per capita terms, physician supply will rise slowly over the next ten years, after which it will decline as equilibrium is reached between the number of trainees and retirees in the face of a growing population. Over this same period of time, the economy will continue to expand and the portion of the economy devoted to health care will rise. At the rate projected, health care expenditures will represent 17% of the GDP in 2020. However, under current training conditions, the supply of physicians will not increase proportionately. Even the addition of larger numbers of NPCs with increased practice prerogatives will fail to meet the need. Indeed, the gap between supply and demand will progressively widen in the years after 2010.

IV. CONCLUSIONS

Studies of the physician workforce face many of the same dilemmas that were encountered in the past. What diseases and treatment modalities will exist in the future? What volume of service will they generate? How will that needed care be financed? Who will provide the care? And how much effort will providers commit to the process? Most importantly, how strong will our economy be and what portion of the national wealth will be devoted to health care services? All of these considerations must be woven into models that set out to define the future requirements for physician services. The Trend Model attempts to do so by incorporating the major dynamics that have affected physician supply and utilization.

The Model predicts a physician shortage beginning in 10-15 years and increasing thereafter. This projection is made at a time when educators and practitioners are confronting an abundant supply of physicians in the face of constrained fiscal resources. It is not easy to plan for winter while in the heat of summer, or to contemplate recession in the midst of prosperity. But both are necessary. So, too, is it important to recognize that powerful dynamics that span decades have led to the conclusion that, within the next two decades, this nation will confront a shortage of physicians in relation to the potentials of medical care, the desires of the public and the capacity of the economy. While the long duration of this projection insulates current educators and planners, it is incumbent on them to begin now to prepare for the needs of tomorrow (11). The Trend Model is offered as a means of defining the magnitude of these future needs.

BIBLIOGRAPHY

1. Lee RI, Jones LW. *The Fundamentals of Good Medical Care*. University of Chicago Press, Chicago, IL. 1933.
2. Graduate Medical Education National Advisory Committee. *Report of the Graduate Medical Education National Advisory Committee to the Secretary, Department of Health and Human Services.* Washington DC. US Department of Health and Human Services. DHHS Publication No. (HRA) 81-652. 1980.
3. Council on Graduate Medical Education. *Third Report.* Washington DC. US Department of Health and Human Services. 1992.
4. Council on Graduate Medical Education. *Fourth Report.* Washington DC. US Department of Health and Human Services. 1994.
5. Weiner JP. Forecasting the effects of health reform on US physician workforce requirement: evidence from HMO staffing patterns. *JAMA*. 272:222-230, 1994.
6. Getzen TE. Population aging and the growth of health expenditures. *J Gerontology: Social Sciences.* 47-S98-104, 1992.
7. Cooper RA, Laud P, Dietrich CL. Current and projected workforce of nonphysician clinicians. *JAMA*. 280:788-794, 1998.
8. Cooper RA, Henderson T, Dietrich CL. Poles of nonphysician clinicians as autonomous providers of patient care. *JAMA*. 280:795-802, 1998.
9. Cooper RA. Perspectives on the physician workforce to the year 2020. *JAMA*. 274:1534-1543, 1995.
10. Cooper RA. Seeking a balanced physician workforce for the 21st century. *JAMA*. 272:680-687, 1994.
11. Reinhardt UE. Academic medicine's financial accountability and responsibility. *JAMA. 284:1136-38, 2000.*
12. Getzen TE. Health care is an individual necessity and a national luxury: Applying multilevel decision models to the analysis of health care expenditures. *J Health Economics*. 19:259-270, 2000.
13. Getzen TE. Forecasting health expenditures: short, medium, and long (long) term. *J Health Care Finance*. 26(3):56-72, 2000.
14. Fuchs V. *Who Shall Live: Health Economics and Social Choice.* World Scientific Publishing Co. Singapore. 1998.
15. Arrow KJ. Uncertainty and the welfare economics of medical care. *Amer Economic Review*. 53:941-973, 1963.
16. Seale JR. A general theory of national expenditure on medical care. *Lancet*. 2:555-559, 1959.

POLICY FORECASTING AND DATA ISSUES

Chair: Clifford Woodruff III
Bureau of Economic Analysis, U.S. Department of Commerce

Assessing the Impact of Government Legislation on BSE in the U.K.,
Sandy D. Balkin, Ernt & Young LLP

The Accuracy of Recent Short-Term Employment Forecasts Obtained by Employer Surveys: The State of the Illinois Experience,
Roy L. Pearson, College of William and Mary
George W. Putnam, Illinois Department of Employment Security
Waleed K. Almousa, Illinois Department of Employment Security

Data Obsolescence and Forecasting,
Othmar W. Winkler, Georgetown University

ASSESSING THE IMPACT OF GOVERNMENT LEGISLATION ON BSE IN THE U.K.*

Sandy D. Balkin
Ernst & Young LLP
1225 Connecticut Avenue, NW
Washington, DC 20036
e-mail: sbalkin@email.com

1 Introduction

Bovine Spongiform Encephalopathy (BSE), also known as Mad Cow Disease, is a disease that has afflicted cows in the U.K. for over a decade. It was first identified in Great Britain in November 1986 by pathologists examining the brains from two cows. There is strong evidence that meat tainted with BSE can cause Creutzfeldt-Jakob Disease, a particular form of a human prion disease characterized by forgetfulness, jerky movements and chronic dementia. The origin of the disease in cattle was not clear, though theories on its spread often focused on supplementary feed containing contaminated meat and bone meal derived from cattle and sheep. As a result of the BSE scare, many countries banned the import of cattle beef from the U.K. In response to economic pressures and to prevent further spread of the disease to humans, the U.K. government introduced various legislative measures.

The U.K. government passed many legislative measures and amendments with three goals: 1) to eradicate the disease by preventing its spread to cattle, 2) to protect public health and 3) to prevent transmission to other animal species. In this study, we investigate the success of initiative (1) by looking at the effect government legislation aimed at preventing spread of the disease had on the the number of confirmed cases of afflicted cows. Specifically, we are interested in the effect of the following on the spread of the disease as given by the Ministry of Agriculture, Fisheries and Food (1996):

1. *The Bovine Spongiform Encephalopathy Order 1988 (SI 1988 No 1039) GB This Order, applicable in Great Britain, was made on 14 June 1988 and came into effect on 21 June (other than the feed ban in article 7 which came into effect on 18 July). It made BSE notifiable and provided for the isolation of BSE suspects when calving. It also introduced a ban on the use of ruminant-derived protein in ruminant feedstuffs with effect from 18 July. The ban was to apply until 31 December 1988 while a review of the rendering processes was conducted. It was introduced as soon as the feed-borne hypothesis had been established in order to prevent further transmission of the infective agent by this route. The primary aim of this measure was the protection of animal health.*
2. *The Bovine Spongiform Encephalopathy Compensation Order 1990 (SI 1990 No 222) GB This came into effect on 14 February 1990. It introduced 100% compensation up to a ceiling for all animals slaughtered under the compulsory slaughter scheme. Its purpose was to support the slaughter policy for the protection of animal health and by compensating owners of affected cattle more realistically for their loss so as to ensure the reporting of suspect cases.*
3. *The Bovine Spongiform Encephalopathy Order 1991 (SI 1991 No 2246) GB This came into effect on 6 November 1991. It consolidated existing BSE legislation and introduced new provisions to prevent meat and bone meal produced from specified bovine offals being used as fertilizer. This was a precautionary measure primarily aimed at the protection of animal health, through grazing of fertilized fields by ruminants.*

The standard technique for determining the effect of such measures is intervention analysis based on ARIMA modeling. Such methods use differencing to remove trends and seasonality from the series prior to analysis. In the past few years, new methods of dealing with trend and seasonal components have been developed allowing these components to be better identified and studied. Here, we examine the impact of government legislation on the observed occurrence of BSE in U.K. cattle using several such techniques. Methods such as traditional ARIMA, structural, and dynamic linear modeling require deterministic inputs from the user specifying the dates when "interventions" or changes in regime occur; by comparison, in automatic ARIMA and Bayesian modeling the method signals where changes occur and the user may then investigate why a change could have occurred. Thus, the aims of this paper are twofold: 1) to

*The views expressed in this report represent the opinions of the author and not necessarily those of Ernst & Young LLP.

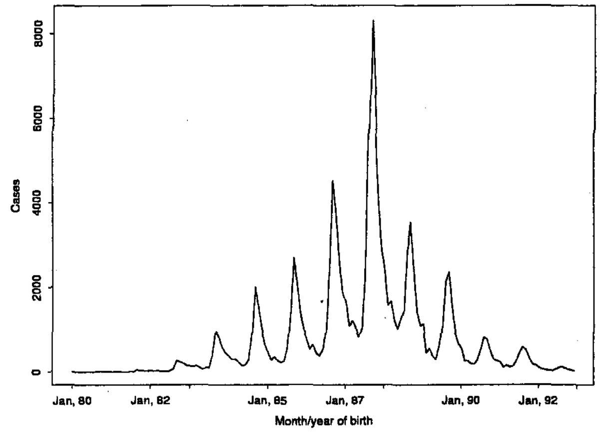


Figure 1: Birth Series

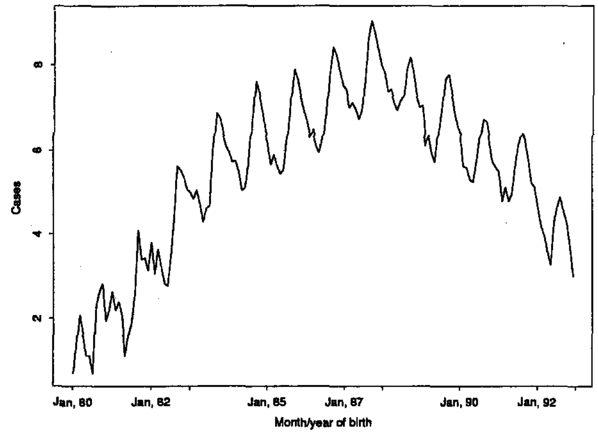


Figure 2: Log Transformed Birth Series

establish the impact of the three legislative acts on the disease and to identify other changes in the structure of the occurrence of the disease and 2) to contrast the various methods used to identify and determine extent of impact.

The data used to perform this evaluation is the number of confirmed BSE cases with known dates of birth aggregated into months of birth from January 1980 to December 1992 as reported by November 1, 1996. The plot of the birth series, shown in Figure 1, exhibits multiplicative seasonality, thus we will use the log transformed series for our analysis. Shown in Figure 2, applying the transformation allows us to better visualize the seasonal regularities and apparently stable variance.

The paper is organized as follows. In Section 2 we present some conventional non-Bayesian models and their corresponding analyses; Bayesian models and their corresponding analyses are given in in Section 3. Section 4 presents a discussion of the time series analysis techniques and how each assessed the effects of the interventions, with the conclusions given in Section 5.

2 Conventional / Non-Bayesian Methods

2.1 ARIMA Modeling

The Box-Jenkins method is a well-known paradigm used to identify the moving average, autoregressive and seasonal components of a stationary time series. In general, when allowing for the series to be transformed and differenced, the Box-Jenkins method provides guidelines to follow when choosing the parameters to identify a model of the form:

$$\phi(B)\Phi(B^s)\nabla^d\nabla_s^D(Y_t - c) = \theta(B)\Theta(B^S)\varepsilon_t$$

The experimenter identifies several possible models and then chooses which is best based upon a set of diagnostics. Forecasts are then based on the selected model. A more detailed explanation of the concepts just presented can be found in Box, Jenkins and Reinsel (1994).

The framework used for evaluating the effect from M interventions is given by

$$Y_t = c + \sum_{i=1}^{M} \frac{\omega_i(B)B^{b_i}}{\delta_i(B)} X_{i,t} + N_t,$$

where c is a constant, $\frac{\omega(B)B^b}{\delta(B)}$ is an impulse response function, X_t is a deterministic variable and N_t follows an ARIMA process as outlined above (Pankratz, 1991). Since legislation primarily institutes permanent changes, we consider step interventions. For a step intervention at time $t = i$, we define

$$X_t = \begin{cases} 0 & t < i, \\ 1 & t \geq i. \end{cases} \tag{1}$$

We establish the interventions at July 1988 ($t = 103$), February 1990 ($t = 122$) and November 1991 ($t = 143$). Going through the Box-Jenkins paradigm of model selection results in the choosing of an

$ARIMA(1,1,0)(0,1,1)_{12}$ model for N_t. The parameter estimates, calculated using *SAS* 6.12, are given in Table 1.

Model Parameter	*Coef.*	*Std. Error*	*t-stat*	*p-value*
MA, Lag 12	0.218	0.090	2.40	0.0174
AR, Lag 1	-0.522	0.076	-6.88	0.0001
Int 1: July 1988	-0.688	0.185	-3.71	0.0003
Int 2: Feb. 1990	-0.138	0.186	-0.74	0.4580
Int 3: Nov. 1991	0.031	0.187	0.16	0.8698
Model Variance	0.072			

Table 1: ARIMA parameter estimates

The mean reduction from Intervention 1 (feed ban) in July 1988 is -0.688 on the log scale. Since $\exp(-0.688) = 0.50258$, the estimated average effect of the legislation amounts to approximately a 50% reduction in the occurrence of the disease. From the t-statistic and its corresponding p-value, we see that this intervention is highly significant. Intervention 2 (compensation act), though statistically insignificant, resulted in a further reduction of approximately 13%, while the third intervention (law consolidation) had little or no effect in curbing the disease.

2.2 Automatic ARIMA Modeling

Automatic ARIMA Modeling is performed via a forecasting package that automatically runs a bank of statistical tests on a series to determine transformations, differencing, lag structure and interventions. For this study, we used Autobox 4.0 by Automatic Forecasting Systems. This software package automates the Box-Jenkins paradigm described in the above section.

Since the primary purpose of this study is to find shifts in the level of the series and not forecasting, we analyze the seasonally differenced series, looking for changes from 12 months prior. Autobox determines that the appropriate model is an AR(2) with level shifts given in Table 2.

Time	*Coef.*	*t-stat*	*% Change*
July 1988	-0.983	-5.28	-62.5%
May 1992	-0.552	-2.73	-42.4%

Table 2: Autobox Level Shifts

Thus, Automatic ARIMA modeling detects Intervention 1 (feed ban) as well as an additional shift in May 1992 that does not directly correspond with any specific legislative act. Figure 3 is a plot of the series with lines drawn at the detected level shifts.

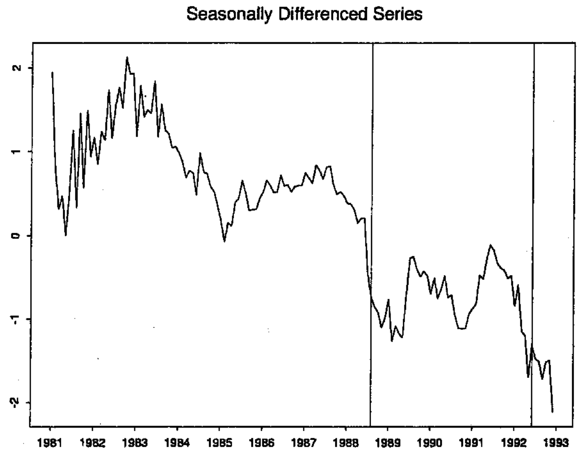


Figure 3: Seasonally differenced Birth Series with level shifts

2.3 Structural Modeling

A univariate structural time series model is formulated in terms of components that have a direct interpretation. A comprehensive theory of structural models is given in Harvey (1989) and demonstrated in Harvey and Todd (1983) and Harvey and Durbin (1986).

Let Y_t be the observed variable. The basic structure model has the form

$$Y_t = \mu_t + \gamma_t + \varepsilon_t, \quad t = 1, \ldots, T, \tag{2}$$

where μ_t, γ_t, and ε_t are trend, seasonal and irregular components, respectively. The process generating the trend is given by

$$\begin{aligned} \mu_t &= \mu_{t-1} + \beta_{t-1} + \eta_t, \quad t = 1, \ldots, T \quad \eta_t \sim NID(0, \sigma_\eta^2) \\ \beta_t &= \beta_{t-1} + \xi_t, \quad t = 1, \ldots, T \quad \xi_t \sim NID(0, \sigma_\xi^2) \end{aligned}$$

The model for a deterministic seasonal pattern is based on a set of trigonometric terms at the seasonal frequencies. So, the seasonal effect at time t is

$$\gamma_t = \sum_{j=1}^{(s-2)/2} (\gamma_j \cos \lambda_j t + \gamma_j^* \sin \lambda_j t) + \gamma_{s/2} \cos \lambda_{s/2} t,$$

where s is the number of seasons in the year and γ_j and γ_j^* are estimated by OLS. All the disturbance terms are

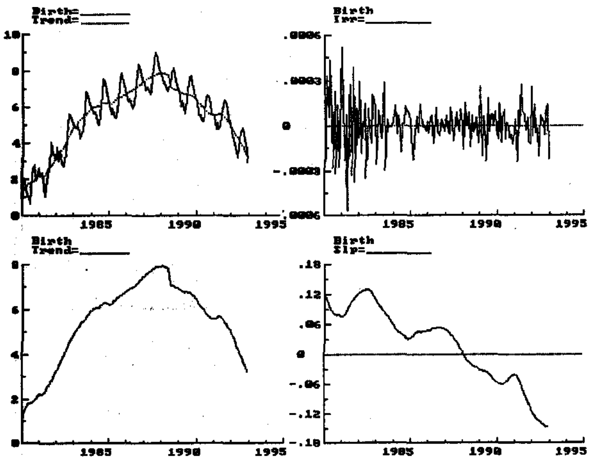


Figure 4: *Structural Components*

Intervention	*Coef.*	*R.m.s.e.*	*t-stat*	*p-value*
Int. 1: July 1988	-0.699	0.183	-3.81	0.0002
Int. 2: Feb. 1990	-0.112	0.183	-0.61	0.5413
Int. 3: Nov. 1991	-0.031	0.185	-0.167	0.8678

Table 3: Intervention Effects

independent of each other as well as the irregular component $\varepsilon_t \sim NID(0, \sigma^2)$. Estimation of the model parameters can be computed in the time domain via maximum likelihood based on the state space representation.

The basic structural model given in equation (2) can be extended to include an instantaneous and constant intervention variable as

$$Y_t = \mu_t + \gamma_t + \lambda X_t + \varepsilon_t, \tag{3}$$

where X_t is defined as in (1). Once the structural time series model is specified, it is put into state space form and fit using the Kalman filter.

From Figure 2, there is evidently a seasonal pattern as well as a trend. Thus, the fitted model is chosen to include stochastic level, slope and trigonometric seasonal components along with the three step intervention variables described in the introduction. A graphical decomposition produced using *STAMP 5.0* is shown in Figure 4. An analysis of the final state of the components gives the coefficients for the interventions as shown in Table 3. The results are similar to those obtained using the ARIMA modeling paradigm. Intervention 1 is the only one of the of the three that appears to have a significant impact on the occurrence of the disease. We do gain some information regarding the seasonal effects on the disease, but those are not of direct interest in this study.

The structural modeling approach also gives us a method for detecting structural breaks based on the auxiliary residuals (Harvey & Koopman, 1992). Fitting a structural model as before but without the interventions results in the auxiliary residual and frequency distribution plots shown in Figure 5. Statistically significant residual values are noted in Table 4. Thus, not only is

Period	*Value*	*p-value*
June 1988	-2.0898	0.0191
July 1988	-3.1630	0.0009
March 1991	2.2073	0.0144

Table 4: Auxiliary Residuals

a structural change apparent at the time Intervention 1 was made active, but it also had a significant effect the previous month, which happens to be the month when the legislation was passed. The significant auxiliary residual value from March 1991 does not correspond to any direct legislation, but is highly positive signaling that the number of infected cows born on March 1991 is unusually high. Further investigation into why this may be the case would be in order to determine possible reasons for this anomaly.

3 Bayesian Methods

3.1 Dynamic Linear Model

Bayesian dynamic linear models (DLM), as explained in West and Harrison (1997) and Pole, West and Harrison (1994) and implemented in Splus in Harrison and Reed (1996), operate according to the principle of *Management by Exception* where an exception is relevant expert information from an external source or a monitoring signal indicating that the performance of the current model is inadequate. DLMs are similar to structural models in that they are specified according to components of interest and use Bayes' Theorem to "learn." By quantifying and using the existing state of knowledge as prior inputs and then combining with observed data quantified probabilistically. The result is the posterior distribution which is used, in general, to specify future beliefs or forecasts.

This sequential model development allows the incorporation of external subjective information concerning future beliefs. For example, in the situation currently being investigated, suppose it was known that approximately 50% of the cows were infected from ruminant-derived protein in ruminant feedstuffs and it was known when Intervention 1 was to be enacted. The investigator can incorporate his prior knowledge into the forecasting model by decreasing the mean level of the series by 50% while increasing the variance to account for the uncertainty surrounding the effectiveness of the legislation.

For this type of analysis, the DLM method includes a tool called Retrospective Assessment. Retrospection is useful in determining "What Happened" given all current information. We will use this type of analysis to assess the impact of government legislation together with the automatic monitoring of model adequacy as detailed in West and Harrison (1997). A linear growth/seasonal discount DLM is applied as in Harrison and Reed (1996) and Cooper and Harrison (1997), with the following prior settings:

Trend	Growth component	
	level:	mean=1.8; se=0.5; disc.=0.95
	growth:	mean=0; se=0.2
Seasonal	Full seasonal	peak/trough
	peak=9; trough=5	mean diff=2; se=1.414; disc.=0.95
Variance	Discount	obsn se=0.2; dof=1; disc.=0.99

The government legislations are incorporated into the model as forward interventions using the following changes:

July 1988		
level	mean=7.6	se=0.4
growth	mean=-0.04	se=0.03
sin1 / sin2	mean=unchanged	se=0.3
cos1 / cos2	mean=unchanged	se=0.3
Feb. 1990		
level	mean=unchanged	se=0.3
growth	mean=unchanged	se=0.03
Nov. 1991		
level	mean=unchanged	se=0.3
growth	mean=unchanged	se=0.03

Intervention 1, the feed ban, includes changes to the means and standard errors of the level and growth components. The other interventions only include increases to the standard error of the two components to reflect the uncertainty associated with the effect those legislative acts will have.

We can see from the retrospective forecast plot, shown in Figure 6, that our model fits the data well. From the level component plot in Figure 7, we see a large drop in July 1998 (Intervention 1), a smaller fall off in February 1990 (Intervention 2), and little in the way of structural

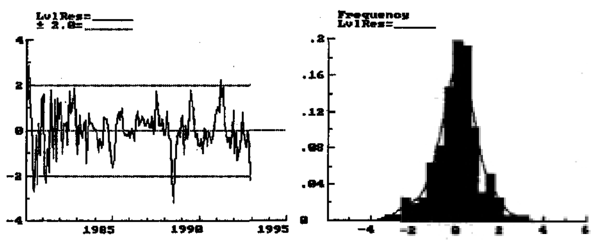


Figure 5: Auxiliary Residual and Frequency Distribution

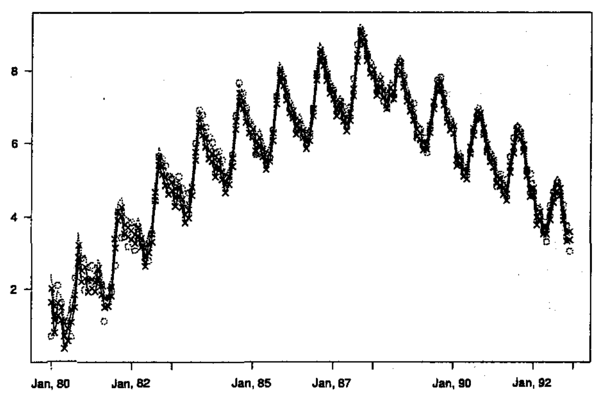


Figure 6: DLM Retrospective Analysis

change from Intervention 3. From Figure 8, the plot of the growth component, we see a large drop at the times of the first and third legislation and a small decrease at the time of the second. This indicates that the feed ban effectively decreases both the level and growth rate of the disease occurrence, while Intervention 3 decreases the growth rate. Figure 9 shows that the amplitude of the seasonal component diminishes over time. Also note that there were no automatic monitoring signals generated after November 1986 while fitting this model implying that, with the inclusion of the three interventions, at no time was the model judged inadequate.

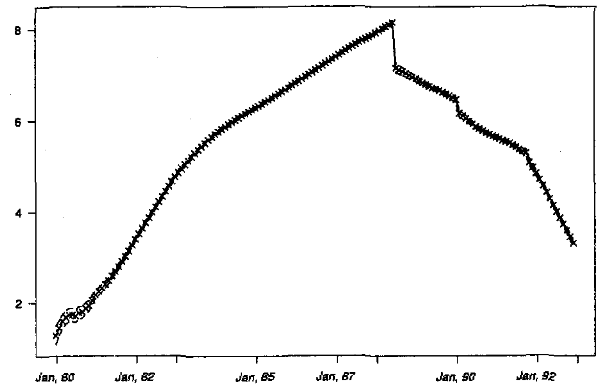


Figure 7: DLM Level Component

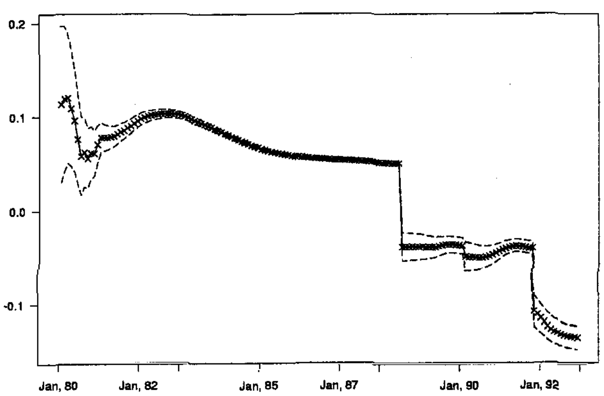


Figure 8: DLM Growth Component

3.2 Gibbs Sampling

The Gibbs sampler is a Monte Carlo method useful in extracting marginal distributions from full conditional distributions when the joint distribution is difficult to integrate. The underlying premise of the sampler is that random realizations can be drawn from the conditional distributions which the sampler can use to provide consistent estimates of the marginal distributions of interest.

For example, consider the case of three parameters $(\theta_1, \theta_2, \theta_3)$ where we are able to draw samples from the three full conditional posterior distributions:

$$f_1(\theta_1|\theta_2, \theta_3, y), \quad f_2(\theta_2|\theta_3, \theta_1, y), \quad f_3(\theta_3|\theta_1, \theta_2, y) \qquad (4)$$

where y is a univariate time series observed at equally-spaced time intervals. The Gibbs sample then works as follows:

1. Consider an arbitrary set of starting values for the three parameters, say $\boldsymbol{\theta}_0 = (\theta_{10}, \theta_{20}, \theta_{30})'$.
2. For "burn-in," generate M sets of random observations drawn iteratively and recursively from the full conditional posterior distributions in (4). Specifically, the first set of random observations $\boldsymbol{\theta}_1 = (\theta_{11}, \theta_{21}, \theta_{31})'$ is obtained as follows

$$\begin{aligned} &\theta_{11} \text{ is drawn from } f_1(\theta_1|\theta_{20}, \theta_{30}, y) \\ &\theta_{21} \text{ is drawn from } f_2(\theta_2|\theta_{30}, \theta_{11}, y) \\ &\theta_{31} \text{ is drawn from } f_3(\theta_3|\theta_{11}, \theta_{21}, y). \end{aligned}$$

3. Generate further sets of random observations, say $\boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_N$, as in the previous step to form a random sample of size N for the parameters.
4. Estimate the posterior marginals from the random sample.

In this study, we use the Gibbs sampler to fit a random level-shift model as described in McCulloch and Tsay

(1993). A time series y_t follows a random level-shift autoregressive model if it satisfies:

$$y_t = \mu_t + x_t, \qquad \mu_t = \mu_{t-1} + \delta_t \beta_t, \qquad x_t = \sum_{i=1}^{p} \phi_i x_{t-i} + a_t,$$

where the δ_ts are i.i.d. Bernoulli random variates such that $\Pr\{\delta_t = 1\} = \varepsilon$, the β_ts are random variates from a given distribution, the ϕ_is satisfy the stationarity condition of the time series x_t and the a_ts are $NID(0, \sigma_a^2)$, all as given in McCulloch and Tsay (1993, 1994). The prior distributions used are given as

$$\begin{aligned} \phi &\sim \mathrm{N}(\phi_0, \mathbf{A}^{-1}) \\ \frac{v\lambda}{\sigma_a^2} &\sim \chi_v^2 \\ \varepsilon &\sim \mathrm{Beta}(\gamma_1, \gamma_2) \\ \beta_t &\sim NID(0, \xi^2) \end{aligned}$$

with hyperparameters ϕ_0, $\mathbf{A}$, λ, v, γ_1, γ_2 and ξ^2 all assumed as known.

Here, the Gibbs sampler is applied to the seasonally differenced series with an $AR(2)$ component as suggested by the PACF plot of the data with hyperparameters fixed at

$$\phi_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \mathbf{A}^{-1} = \begin{bmatrix} 73 & 72 \\ 72 & 73 \end{bmatrix}$$

where $\mathbf{A}^2$ is the correlation matrix between AR coefficients and $v = 2$. The other hyperparameters are determined based on the residual variance of fitting an $AR(2)$ model to the data. The only "user" input is the prior belief probability that a level shift occurs. We set $\gamma_1 = 1$ and $\gamma_2 = 99$ to reflect a prior belief that a given level shift occurs with probability 0.01.

The Gibbs sampler was iterated for 10,000 iterations with the first 4,000 as the burn-in sample. Figures 10 show the estimated mean process μ_t with one-standard-error limits of μ_t and associated posterior probability of shifts for the three prior beliefs of ε.

Major level shifts and their posteriors were obtained for $\varepsilon = 0.01$ from Figure 10 and are given in Table 5. Note that since we are primarily concerned with assessing the impact of legislation on the occurrence of the disease, and since the disease was only first identified in November 1986, we only analyze level shifts after that date.

The drop in July 1988 can be attributed to Intervention 1, the ban on use of ruminant-derived protein in ruminant feedstuffs. The mean reduction in this month is -0.939 on the log scale. Since $\exp(-0.939) = 0.39$, the estimated average effect of this ban on the occurrence of BSE amounts to approximately a 61% reduction in the occurrence of the disease.

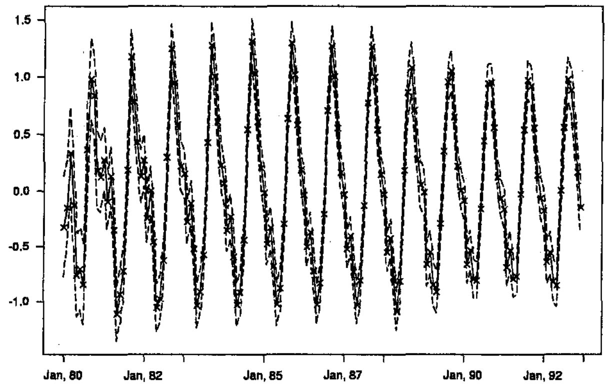


Figure 9: DLM Seasonal Component

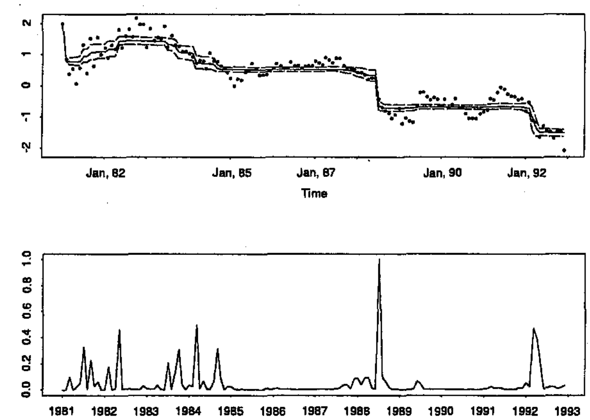


Figure 10: Gibbs Sampler: $\gamma_1 = 1$, $\gamma_2 = 99$

Time	July 1988	March 1992	April 1992
Probability	0.984	0.574	0.326
Shift	-0.939	-0.436	-0.231

Table 5: Major Level Shifts and their Posterior Probabilities, *(Post Nov. 1986)*

A further reduction appears to have occurred during March and April 1992. The total mean reduction during this time period is -0.667 on the log scale resulting in estimated average reduction of 49% from the previous level. Apparently, however, no specific government legislation was passed around this time. Cooper and Harrison (1997) offer as a possible explanation that cattle born after January 1992 were thought to have been carefully protected from infection via the main sources and that the infection is converging to a new level, possibly determined by a non-feed-based source of infection.

4 Discussion of Results

This study looks at the effect of government legislation on the occurrence of BSE in the U.K. using a variety of time series analysis techniques. Each method has its own strengths in this type of study. Each method recognized Intervention 1, the ban on the use of ruminant-derived protein in ruminant feedstuffs, as a highly significant structural change in the series. ARIMA and structural modeling gave similar results when analyzing the effects of the three interventions, but structural modeling gave additional information via the auxiliary residuals and is capable of providing considerable information about the seasonal patterns of the series, if so desired. The Automatic ARIMA modeling, the auxiliary residuals from the structural model and the Gibbs Sampler both signaled some additional structural breaks which did not directly correspond to legislation. This extra structure is most likely caused by the 5 year incubation period of the disease in cattle. The structural and DLM methods allow the model parameters to change over time. The DLM method goes even further in that at each new observation, the model is checked for inadequacies via automatic monitoring. Also, the structural and DLM methods provide information regarding the growth rate of the disease in addition to its level.

Each method has drawbacks as well. For ARIMA modeling, there are at least four approaches for model identification when including interventions (Kendall & Ord, 1990). Each method has its attractions based on the behavior of the individual intervention. Also, it is traditional in ARIMA modeling to remove trend and seasonal dependencies via differencing and transforming. In this case, a seasonal and a regular difference are taken to induce stationarity. This may influence the effect of the interventions. As an alternative, instead of removing dependencies, we can incorporate them into the model as in structural modeling and dynamic linear modeling. These methods model level, trend and seasonal structure as unobserved components instead of removing their effects. Thus, using either of these methods provides additional information which may explain some of the observed fluctuations. However, when implementing the DLM method of forward interventions, the user is required to have expert knowledge of the effect of the intervention, which is usually difficult to obtain without looking at future values of the series. Otherwise, as in this case, all the experimenter can do is increase the uncertainty level associated with the future observations. The Gibbs sampling approach has similar drawbacks in that the results are very dependent on the quality of prior information. Also, a seasonal difference was necessary for the Automatic ARIMA and Gibbs Sampling methods which effectively changes the model to look for changes in level from time t from the level 12 months prior.

5 Conclusion

The legislative acts examined in this study were designed to protect cattle and prevent spread of the disease. The hypothesis that BSE was mainly being spread through ruminant feed seems quite plausible in that the ban on the feed drastically reduced disease occurrence. This measure had no effect on cattle infected before its introduction and its effectiveness may have taken some time due to noncompliance by cattle raisers and the 5-year incubation period of the disease.

Based on the analysis provided by these methods, the feed ban resulted in an approximately 50% reduction in the disease, with a 61% decrease from the year prior to the introduction of the ban. The compensation and consolidation acts did not necessarily directly affect the level of the disease, but each did have an impact by reducing the infection rate. Finally, it appears that, due to the 5-year incubation period, the series reaches a new level approximately 60 months after the implementation of the feed ban; this new level is possibly determined by a non-feed-based source of infection.

References

Cooper, J. D., & Harrison, P. J. (1997). Modelling the Observed Bovine Spongiform Encephalopathy Epidemic: A Bayesian Approach..

Donnelly, C. A., Gore, S. M., Curnow, R. N., & Wilesmith, J. W. (1997). The Bovine Spongiform Encephalopathy Maternal Cohort Study: Its Purpose and Findings. *Journal of the Royal Statistical Society (Series C)*, *46*(3), 299–304.

Donnelly, C. (2000). Statistics Under the Public Gaze: BSE and vCJD. *RSS News*, *27*(9).

Donnelly, C. A., & Ferguson, N. M. (2000). *Statistical Aspects of BSE and vCJD.* Chapman & Hall / CRC.

Harrison, P. J., & Reed, R. J. (1996). Dynamic Linear Modelling with S-Plus. *Research Report 310.* Department of Statistics, University of Warwick.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge University Press.

Harvey, A. C. (1991). *Time Series Models.* The MIT Press.

Harvey, A. C., & Durbin, J. (1986). The Effects of Seat Belt Legislation on British Road Casualties: A Case Study in Structural Time Series Modelling. *Journal of the Royal Statistical Society, Series A*, *149*, 187–227.

Harvey, A. C., & Koopman, S. J. (1992). Diagnostic Checking of Unobserved Components Time Series Models. *Journal of Business & Economic Statistics*, *10*(4).

Harvey, A. C., & Todd, P. (1983). Forecasting Economic Time Series with Structural and Box-Jenkins Models. *Journal of Business & Economic Statistics*, *1*, 299–315.

Kendall, M., & Ord, J. K. (1990). *Time Series* (Third edition). Edward Arnold.

Koopman, S. J., Harvey, A. C., Doornick, J. A., & Shephard, N. (1995). *Stamp 5.0: Structural Time Series Analysis Modeller and Predictor.* Chapman & Hall.

McCulloch, R. E., & Tsay, R. S. (1993). Bayesian Inference and Prediction for Mean and Variance Shifts in Autoregressive Time Series. *Journal of the American Statistical Association*, *88*(423), 968–978.

McCulloch, R. E., & Tsay, R. S. (1994). Bayesian Analysis of Autoregressive Time Series via the Gibbs Sampler. *Journal of Time Series Analysis*, *15*(2), 235–250.

Ministry of Agriculture, Fisheries and Food (1996). Programme to Eradicate BSE in the United Kingdom. *Ministry of Agriculture, Fisheries and Food.*

Pankratz, A. (1991). *Forecasting with dynamic regression models.* Wiley.

Pole, A., West, M., & Harrison, J. (1994). *Applied Bayesian Forecasting and Time Series Analysis.* Chapman & Hall.

Tiao, G. C., & Tsay, R. S. (1994). Some Advances in Non-linear and Adaptive Modelling in Time-Series. *Journal of Forecasting*, *13*, 109–131.

West, M., & Harrison, J. (1997). *Bayesian Forecasting and Dynamic Models* (Second edition). Springer-Verlag.

The Accuracy of Recent Short-Term Employment Forecasts Obtained by Employer Surveys: The State of Illinois Experience

Roy L. Pearson, College of William & Mary
George W. Putnam, Illinois Department of Employment Security
Waleed K. Almousa, Illinois Department of Employment Security

The U. S. Workforce Investment Act of 1998 and generation-low U.S. unemployment rates have made short-term industry employment and occupational forecasts at the regional level an immediate priority. Employer surveys of future employment are one forecasting approach being tested.

A necessary condition for survey-based forecasting to be efficient is that employers can predict their future employment with reasonable accuracy. The minimum standard for accuracy is that the forecasts are more accurate than a naïve, no-change forecast. If naïve forecasts are as accurate as survey-based forecasts, then currently available employment and occupational information is the more efficient basis for short-term planning. Therefore the accuracy of no-change forecasts is a meaningful benchmark for gauging the relative accuracy of the survey forecasts.

The database analyzed here is 15,847 three-month-ahead forecasts by 13,025 different establishments obtained from quarterly surveys of Illinois employers during 14 quarters from 1995.4 through 1999.1. The quarterly survey solicits information for the last month of each quarter and a forecast for the last month of the next quarter. For example, employers receive the first-quarter survey in either the second or third week of March. The respondent provides an estimate of the establishment employment for the reference month of March and a forecast for June employment. In the second-quarter survey, the reference month is June and the forecast month is September. The surveys in the third and fourth quarters follow a similar solicitation pattern. Therefore the quarterly-solicited forecasts all are monthly forecasts for three months in the future.

Each quarterly survey was mailed to a random sample stratified by industry of approximately 3,400 establishments selected from the Illinois ES202 database. The lowest quarterly response rate was 38.5% and the highest, 63.3%. The mean rate was 47.2% and the median was only slightly lower (46.5%). Nine of the fourteen quarterly response rates fell within the narrow range of 45% to 49%. The responses then were edited for obvious errors using procedures based only on the available information at the time of the survey, not the actual ES202 data received subsequently. Thus the editing procedures can be applied in the future at the time the survey responses are received.

The final edited database of 15,847 observations is a monthly average of 1,132 establishments with 42,836 employees. For the 14 survey months, the reported employment averaged 0.74% of the total Illinois employment, with a median of 0.71% and a standard deviation of 0.16%. The highest percent was 1.04% in 1995.4; and the lowest, 0.49% in 1996.4.

The first question examined here is the accuracy of the surveys in predicting total Illinois employment three months in the future, based on the actual ES202 reported employment. Subsequently, the forecasting accuracy of individual establishments is assessed, using the individual firm records as observations.

Survey Predictions of Total Employment

The sample respondents' estimates of current and three-month-ahead employment provide the key piece of information, the predicted percent change, necessary to forecast the percent change in the Illinois total ES202 employment three months in the future. A preliminary test revealed that a predicted growth rate based on the sample's total current-month and predicted-month employment was less accurate than naïve forecasts in predicting total employment levels and growth rates. Two reasons are the response bias in the surveys in distribution of employment by industry sector and by size of firm.

The sample and total employment distributions were derived for nine industry sectors: agriculture; mining; construction; manufacturing; transportation, communication, and public utilities (TCPU); finance, insurance, and real estate (FIRE); services; and government. The average sample percentage distributions for the 14 reference months, in order, were:
0.3, 3.1, 1.8, 28.0, 7.1, 7.1, 3.3, 47.3, and 1.9.
The average population distributions for the 14 reference months in percents were, respectively:
0.8, 0.2, 4.1, 17.1, 5.6, 23.3, 6.8, 28.5, and 13.7.
Mining, manufacturing, TCPU, and services employment were over-represented in the sample, particularly services employment. Agriculture, construction, FIRE, trade, and government employment were under-represented, especially trade and government.

To reduce this response bias, the sample's current and projected month employment levels for each sector were summed, and predicted growth rates calculated for each of the nine sectors. These growth rates then were weighted by the population employment distributions, to derive a predicted growth rate for total employment.

The population distributions used were the ES202 distributions from the same month in the preceding year for the forecasted month. For example, the December 1995 sample forecasts were for March 1996. Therefore the sector growth rates were weighted by the total employment distribution for March of 1995. Using the prior years' forecasted-month distribution, instead of the prior year's reference month, adjusted for seasonal fluctuations in industry sector employment as well as response bias in the sample. This weighting approach can be used in the future, since the ES202 validation and verification process typically is completed within six months, and accurate employment distributions will be available prior to the month the surveys are conducted. For example, the March 1995 ES202 final data were available well before the December 1995 survey, allowing the December 1995 forecasted sector growth rates for March 1996 to be reweighted at the time that the December responses were received.

The sample responses also were biased towards large establishments. For the 14 survey months, the average employment of the survey establishment was 38.3 employees, compared to an average of 18.4 employees in the total population. No explicit attempt was made to adjust for this bias. However, the reweighting by industry sector, particularly reducing the weights for services and manufacturing and increasing the weight for trade employment, substantially reduces this bias.

The predicted growth rates for each of the nine industry sectors for each of the 14 forecasted months were compared to the actual growth rates for those sectors as shown by the subsequent full population ES202 data, to gauge sector accuracy. However, the most important sector variable was not its growth rate or error, but its contribution to the predicted growth rate for total employment. That contribution was calculated for the 14 prediction months by multiplying each sector's forecasted growth rate by its percent of total employment in the corresponding month from the preceding year. The sum of these contributions is the survey's forecast of the percent change in total Illinois employment for the three-month horizon.

The predicted and actual percent changes in total employment are given in Table 1. The percent change error, PCE, by months is calculated throughout this paper as the predicted percent change minus the actual percent change. Thus negative errors show an underestimate of the change; and positive errors, an overestimate.

The mean percent change error, MPCE, for the 14 months is –0.5%, an underestimation bias. The mean absolute percent change error, MAPCE, is 1.3%. If no-change is the forecast, then the mean absolute actual percent change, of 1.6%, is the MAPCE. Therefore the survey results predict the change in total employment somewhat more accurately than a naïve assumption of no-change. How much more accurately also is quantified by Theil's U, the square root of the [(sum of the squared PCE)/(sum of the squared actual percent changes)]. Theil's U provides an index ranging from zero to one of the size of the error in percent change forecasts relative to a no-change benchmark. Thus the Theil's U in this case of 86.8% shows the monthly errors were 13.2% lower on the average than naïve, or status quo, forecast errors.

Note that the percent change error, PCE, and the MAPCE derived from it use the base period as the denominator, while the more commonly used percent error and MAPE are based on ending period values in the denominator. The relatively small monthly percent changes in table 1 make the difference between the MAPCE and MAPE negligible, 0.006%.

The base period is the reference point for the size and direction of predicted growth rates, the key survey-provided information. Therefore error measures such as PCE and MAPCE with base-period denominators, used exclusively in this paper, are more consistent and appropriate accuracy measures in this situation than those based on ending values.

The size of the predicted percent change is not the only useful survey information. The direction of the predicted future change also is important, used extensively by the Illinois Department of Employment Security in reporting survey results and forecasting future employment. Moreover, the public remembers an error in predicting the direction of change — particularly in predicting a downturn that does not occur — more readily than the size of even a large error. Therefore the accuracy of the signals of direction is a relevant question, considered extensively in this paper.

Total Illinois employment declined in each March from the levels in December, and then increased significantly from March to June, indicating a significant seasonal variation. The survey responses correctly predicted decreases in the four March months, and increases in the four June months. In addition to correctly predicting the direction of change for these eight months, the MPCE was 0.2%, a negligible overestimation bias, and the MAPCE was only 1.2%.

However, total Illinois employment rose in each of the three Septembers and three Decembers, but the surveys incorrectly predicted declines in five of those six months. These directional errors also made the bias and size of the errors in those six months larger, with a MPCE of –1.5% and a MAPCE of 1.5% — and worse than a no-change prediction for those six months, for which the MPCE is –0.5% and MAPCE 0.5%.

The very low 0.5% average actual change for these six months means that naïve forecasts will have low errors, making them hard to beat even if the direction of the survey prediction is correct. However, an analysis of the industry sector contributions to the overall growth rate forecast reveals some structural problems that may be reduced in the future. For each of the thirteen months in which the percent change errors in the total growth rate were not zero (to one decimal place), a sector contribution was clearly identifiable as the major source of the overall forecast error, and these are identified in Table 1. As shown previously, the largest sector weights, in descending order, were services, trade, manufacturing, and government. Therefore an error in these sectors contributes more to the total error than one in the remaining five industry sectors. However, among the thirteen months, the services sector was the major source only once; trade, twice; and manufacturing, once. The highly volatile construction sector, with an average absolute percent change of 10.6% but a very small employment share, 4.0%, was the main cause only twice. On the other hand, the government sector, with an average absolute percent change of only 1.1%, was the major error source in six of the thirteen months, almost one-half.

Furthermore, the government sector forecasts were the major cause of the survey's error in four of the five periods in which the survey wrongly predicted that total employment would decline. That realization was a signal to examine the government sector forecasts for these four months more closely. The government sector's survey responses in the four reference months (June and September of 1996 and 1997) were a below-average percent of the population, representing only 0.3% of the government establishments and 0.2% of government employment. Using that sample's projected growth rates for September and December of 1996 and 1997 as being representative of the total government sector employment yielded underestimation errors in all four months, averaging 13.4%.

If the government sector forecasts had been naïve no-change predictions for the six months in which that sector was the major source of error, the 14 percentage change forecasts for total employment would have had a mean percent error of only –0.2%; a lower MAPCE of 1.0%; and a Theil's U of 68.2%, nearly twenty percentage points below the actual U. Also, the signs of the forecasts would have correctly signaled the direction of change for 10 out of 14 months, instead of only 9 out of 14.

Table 1- Overall Monthly Accuracy

Predicted in Month	Predicted for Month	Pred. % Change	Actual % Change	% Error, Pred-Act	Major Source of Error
1995.12	1996.03	-1.3%	-1.7%	0.4%	GOVERNMENT
1996.03	1996.06	1.2%	2.4%	-1.2%	FIRE
1996.06	1996.09	-2.5%	0.2%	-2.7%	GOVERNMENT
1996.09	1996.12	-0.3%	0.7%	-1.0%	GOVERNMENT
1996.12	1997.03	-2.0%	-2.0%	0.0%	None
1997.03	1997.06	2.7%	3.0%	-0.2%	SERVICES
1997.06	1997.09	-0.9%	0.0%	-0.9%	GOVERNMENT
1997.09	1997.12	-2.3%	1.0%	-3.3%	GOVERNMENT
1997.12	1998.03	-4.3%	-1.9%	-2.4%	TRADE
1998.03	1998.06	4.9%	3.0%	1.9%	CONSTRUCTION
1998.06	1998.09	0.2%	0.1%	0.1%	TRADE
1998.09	1998.12	-0.4%	0.7%	-1.2%	CONSTRUCTION
1998.12	1999.03	-0.5%	-2.3%	1.8%	MANUFACTURING
1999.03	1999.06	4.4%	2.9%	1.5%	GOVERNMENT

Naïve Forecast:		Survey Forecast:	
MPCE	0.4%	MPCE	-0.5%
MAPCE	1.6%	MAPCE	1.3%
		Theil's U	86.8%

The forecasts by the government respondents also underestimated the sample's future government sector employment, *in all four of the months where* government was the major cause of the turning-point error, with an average percent change error of –17.7%, pointing to possible problems in reporting and verification. Such an analysis illustrates how an initial review of sector contributions to the overall forecast and forecast error, followed by a closer examination of the forecasts in the sector that most frequently is the major contributor to the overall forecast error, can identify problems that may be ameliorated by revising the sampling or response verification process.

In sum, the Illinois survey-based forecasts of total employment's percentage and direction of change three months in the future have been moderately more accurate than no-change forecasts, a significant achievement in an environment where the average actual changes have been relatively small. Furthermore, the analysis here demonstrates that ongoing monitoring of the forecast results can identify problem areas whose elimination will substantially improve accuracy.

Individual Establishment Predictions

The individual firm analysis serves three primary purposes: To assess forecasting ability at the *establishment level; to determine whether that ability* varies by firm characteristics or time period; and to yield insights into how such a survey process may be improved.

When the objective is to predict total employment growth, a survey establishment's relative impact is based on its number of employees. Here, where the focus is the forecasting accuracy of the individual establishments, each establishment's response carries equal weight, regardless of size – a 'one establishment, one vote' approach. Thus the analysis is based on the 15,847 individual establishment observations and subsets of that total. We present the results first for all establishments, then for subsets based on seasonality and establishment size.

Error Measures

Two types of error measures were selected as appropriate, based on the nature of the data and its uses: Error measures based on the percent change from the beginning to ending months; and direction errors based on the nature, but not the specific size, of the predicted percent changes. Table 2, summarizing the error analysis for the full sample of 15,847 observations, provides a reference for the reporting format of the error measures described below.

The error measures for assessing the size of the percent change error are the mean percent change error, MPCE, and its two components, the mean predicted percent change, MPC, and the mean actual *percent change, MAC. These indicate the average* direction of the predicted and actual percent changes as well as the direction of the average bias.

For size of the percent error, we give the mean absolute percent change error, MAPCE, and compare it to the mean absolute actual percent change, MAAPC. The MAAPC also is the absolute percent error for a naïve no-change forecast. If, on average, MAPCE is less than MAAPC, then the forecasted percent change differs from the actual by less than the actual differs from zero, one indication that the forecast is on average superior to a no-change prediction. Theil's U also is reported, to provide an index of the size of the percent change forecast error relative to that of a no-change benchmark.

For analysis of the accuracy of the direction-of-change signals, we use Theil's Prediction-Realization tables (Theil, 1966). The basic table has nine cells, for pairing the three possible predictions — increase, no-change, or decrease, — with the realized outcomes. The nine cells show the frequencies of the nine possible combinations. Row sums give the distribution of the predictions; and column sums, of the realized outcomes. The diagonal of this table sums to the percent of correct forecasts of direction.

Underneath the table, we repeat the percentages of forecasts with the correct sign; and also show separately the percentages of correct, underestimated, and overestimated percent changes. These two rows are key information for assessing the accuracy of the direction signals.

The percentages of correct, underestimated, and overestimated percent changes are based on the distribution of pairwise outcomes for 13 possible types of pairings, grouped by correct and incorrect signs. These pairings differ conceptually from the mean percent change errors described above. With the percent change error, PCE, any prediction higher on the numerical scale than the actual value is an overestimate, even if both values are negative. Thus if the predicted percent change is –20%, and the actual percent change is –50%, the PCE will be +30%, indicating that the PCE is an overestimate of the "growth" as well as the level of employment.

However, in the 13 directional pairings the position relative to zero is the basis for comparison. In these pairings, the example of a predicted percent change of –20% paired with an actual percent change of –50% is classified as a predicted decrease that is the correct direction but with an underestimate of the extent *of the percent decline. The two different approaches are* not inconsistent, merely different views of the situation. For example, in this illustration, the PCE indicates that the firm overestimated "growth" and the level of future employment by 30%, while the detailed direction of change analysis shows that error in predicting level is due to underestimating the percentage decline in employment. The summary in the prediction-realization analysis showing the underestimation and overestimation percentages is derived with zero as a base for evaluating directional changes, thus providing a different view of the decomposition of forecast bias than provided by the MPC, MAC, and the MPCE.

No-change is both a viable forecast and non-trivial outcome for short-term employment. For the 15,847 observations, 37.1% of the actual outcomes were exactly the same employment in the ending month as in the initial month. Therefore the definition used for a no-change forecast is an exact match of the two monthly employment values.

An additional issue is whether the establishments' forecasts provide additional useful information, versus knowing only the past percentages for the three possible outcomes — increase, no-change, or decrease. For this analysis, Information Gain and Relative Information Gain tables, not shown here, were constructed following the approach in Theil (1966, chapter 12). The relative information gain from obtaining the firms' forecasts is summarized in Theil's Q, a zero-one index of the quality of the directional forecasts. This index is a geometric mean of the relative information gains from the firm's three possible types of predictions when information is available about the past distribution of actual outcomes. If finding out that the establishment is forecasting an increase, no-change, or decrease adds significant information beyond merely knowing from past experience the probable frequencies for outcomes, Theil's Q is closer to one.

The Q is a more comprehensive measure of the forecast quality than the percent correct, since the Q takes all nine prediction-realization cells into account, instead of just the three on the diagonal that are summed to get the percent of correct forecasts. In our observations for individual establishments, no-change is a very frequent forecast, and such forecasts offer less relative information gain than the prediction of an increase or decrease. The Theil's Q adjusts for this difference in relative information gains, generally discounting the no-change forecasting results and placing a premium on the correctness of the increase and decrease forecasts.

Table 2 - All Establishments, All Quarters
Sample Size of 15,847

Error Measures for Three-Month Percent Change

Size of Percent Change	
Mean Predicted % Chg., MPC	2.9%
Mean Actual % Chg., MAC	3.3%
Mean Percent Change Error, MPCE	-0.4%
MAPCE	18.0%
Mean Abs. Actual % Chg., MAAPC	16.9%
Theil's U	1.009
Theil's Q	0.455
(0-1 Index for Quality of Change Prediction)	

Prediction-Realization Table

	Realization			
Prediction	Increase	No Change	Decrease	
Increase	10.8%	4.7%	6.2%	21.7%
No Change	17.2%	29.8%	17.5%	64.5%
Decrease	3.3%	2.7%	7.9%	13.9%
	31.3%	37.1%	31.6%	100.0%

Correct Sign	Increase	No Change	Decrease	
	10.8%	29.8%	7.9%	48.5%

Error by	Correct %	Underest.	Overest.	
Over/Under	32.9%	49.4%	17.7%	100.0%

All Establishments, All Months

The accuracy analysis for the 15,847 observations is presented in Table 2. The mean predicted percent change, MPC, of 2.9% is below the mean actual percent change, MAC, thus the mean percent change error, MPCE is –0.4%, a slight tendency to underestimate the future changes.

The MAPCE of 18% exceeds the MAAPCE, and Theil's U is one. Therefore the error measures based on the size of the predicted percent change indicate that the establishment forecasts on average are no more accurate than a naïve no-change forecast.

As predictors of direction of change, the firms were correct only 48.5% of the time. The majority, 60%, of their correct forecasts came from accurately predicting no-change for the coming period. However, they have a predilection to predict no change far more often than it occurs – 64.5% of their forecasts versus only a 37.1% realization – and their no change forecasts have been wrong over half the time (29.8% correct divided by 64.5% total). Similarly, forecasts of an increase have been wrong slightly more than 50%. The only relatively significant information signal has been in predicting a decrease. For the one in seven instances when they predicted a decrease, they were correct in more than half of the cases. However, the overall result is a Theil's Q of .455, a marginal value. If the firms were proficient in forecasting, we would expect both the percent correct and the Theil's Q to be over 50%.

In sum, the individual forecasts on average are no better than no-change forecasts in predicting the size of the firm's future employment and percent change from current employment. In direction, only 48.5% correct signals is an unreliable record, although it is better than the 37.1% that would be correct if all of their forecasts were no-change. A major weakness in their responses is the no-change prediction nearly two-thirds of the time. If instead about 60% of their predictions were for change, their forecasting accuracy potentially would be higher, especially since decreases are frequent (31.6%) and they do relatively well in anticipating them. Providing them with prior realization distributions, similar to the 31% increases, 37% no-change, and 32% decreases shown here, would give them domain knowledge that may lead them to rely less often on no-change forecasts.

Seasonality and Establishment Accuracy

Predictions of total employment growth were more accurate for March and June months when seasonality clearly was a factor. A corollary hypothesis is individual establishments will be more accurate forecasters when seasonality is part of the employment variations.

The impact of seasonality on individual accuracy was analyzed two ways. First, by comparing the four subsets of months of the year; and secondly, by comparing a subset of firms in very seasonal industries with a subset of establishments in less seasonal industries.

The Four Seasons

The sample sizes by the forecasted months are: March, 4,220; June, 4,608; September, 3,476; and December, 3,543. Our initial hypothesis was that significant and predictable seasonal changes from December to March and March to June would be accompanied by more accurate forecasts. However, the necessary condition for testing this hypothesis was not present in the individual establishment data. Pairwise tests of the sample means for actual employment did not indicate that the differences across any of the months could reasonably be attributed to seasonality. For the full sample of individual respondents to the fourth and first quarter surveys, seasonal variation in the means apparently is overshadowed by random variation or cancelled by offsetting differences across establishments.

Selected error measures from the four monthly analyses are shown in the first four rows of Table 3. The quarterly errors for size of percent change and for direction of change are very similar – equally marginal – across the four forecasted months. The major differences occur in the accuracy of the percent change predictions.

The MAAPC, mean absolute percent change, of actual employment for the individual establishments is larger in the June dataset, for changes from March to June, than in the other three. Moreover, the firms' ability to predict these June changes is relatively higher, as indicated by the mean absolute percent change error only 0.2% higher than the MAAPC and by the Theil's U of .84. However, that relative advantage in predicting the size of the percent error does not carry over to a relatively superior performance in predicting the direction of change. The 48.1% percent correct for direction and Theil's Q of .46 are very similar to the values in the other three quarters. One reason for the similarity in the direction errors is the percent of no-change forecasts for June is 64.0%, not significantly less than for the other three months.

High Seasonality Industries Versus Low Ones

Firms were designated as either seasonal or nonseasonal based on their 2-digit SIC industry classification code and the seasonal factors computed by the Bureau of Labor Statistics for Illinois industries covered in the Current Employment Statistics program. The definition of a seasonal industry has two components: 1) The industry must demonstrate a monthly fluctuation in the level of employment that is replicated in each of the most recent three years, and 2) The range between the maximum and minimum monthly seasonal factors must exceed six percent. The industries that met these criteria, i.e., seasonal, are: general building contractors, heavy construction contractors, special trades contractors, trucking and warehousing, air transportation, transportation services, building materials stores, general merchandise stores, furniture and home furnishing stores, eating and drinking places, real estate, personal services, amusement and recreation services, educational services, and membership organizations. All other industries were designated as nonseasonal.

The seasonal dataset consists of 4,330 observations; and the nonseasonal, 11,517. The full error analysis tables for these two groups are available on request, and selected error measures from those tables are in the last two rows of Table 3.

Our initial hypothesis was that establishments in seasonal industries would be somewhat more accurate forecasters because they would be aware of their monthly seasonal variations and able to predict them with reasonable accuracy. Surprisingly, the data do not support that hypothesis. The seasonal firms do experience larger percent changes in their employment, but their MAPCE also is significantly higher. The net result, as indicated by the Theil's U of 1.02 versus .99 for nonseasonal firms, is the seasonal firms' accuracy relative to a naive forecast is no better than the accuracy of nonseasonal establishments.

One would at least expect seasonal firms to be more accurate in predicting the direction of change, but both the percent correct and Theil's Q show that they do no better than nonseasonal firms. One likely reason is other factors influence direction over three-month spans more than does seasonality. An indication of this possibility is that for percentages of increases and decreases, decreases were 49.4% of the total for seasonal firms, but a somewhat higher 50.5% for nonseasonal firms, implying that decreases are not primarily attributable to seasonality.

Another interesting probable reason for the lack of superior accuracy is that the seasonal firms have predicted no-change exactly the same 64.5% of the time as nonseasonal ones, even though the seasonal firms do experience a somewhat smaller percentage of no-change outcomes. That practice by seasonal establishments makes incorrect no-change forecasts 35.8% of their total predictions, versus 34.3% for nonseasonal firms. That implies that establishments in highly seasonal industries are not effectively recognizing and incorporating their seasonal variations into their employment forecasts.

Table 3 - Error Summary for Seasonality

	For Size of % Change Error				For Direction Error		No-Change Forecasts Versus Reality		
Forecasted Month	MAPCE	MAAPC	MAPCE-MAAPC	Theil's U	% Correct Signal	Theil's Q	% No Chg. Pred.	% No Chg. Real.	Wrong NC Pred.
March	18.6	17.4	1.2	1.00	48.5	0.45	63.5	37.1	34.0
June	20.2	20.0	0.2	0.84	48.1	0.46	64.0	36.2	35.5
September	14.8	13.7	1.1	1.08	49.3	0.43	62.7	37.7	32.5
December	17.7	15.5	2.2	1.52	48.4	0.46	67.9	37.7	36.6
Industry Type									
Seas	22.7	21.9	0.8	1.02	47.8	0.46	64.5	35.2	35.8
Nonseas	16.2	15.0	1.2	0.99	48.8	0.45	64.5	37.8	34.3

Table 4 - Error Summary for Establishment Size

	For Size of % Change Error				For Direction Error		No-Change Forecasts Versus Reality		
Size	MAPCE	MAAPC	MAPCE-MAAPC	Theil's U	% Correct Signal	Theil's Q	% No Chg. Pred.	% No Chg. Real.	Wrong NC Pred.
Small	19.1	18.1	1.0	1.01	49.3	0.46	69.1	40.8	35.7
Large	9.3	8.3	1.0	1.14	50.5	0.37	29.8	16.8	19.2

Small Versus Large Establishments

The differences by establishment size are the most significant and interesting among the 24 subsets we examined. The small firm sample has 13,996 observations; and the large firm sample, 1,851. Table 4 with selected error measures is provided for convenience, along with the full error analyses in Tables 5 and 6.

The small establishment environment is quite different from that of the large ones, and so are the forecasts and their accuracy results. Whether or not the small firms are better forecasters than the large ones is judgmental, depending on which error measures you emphasize and how you interpret the results. However, the fact that they are different is indisputable, so we describe those differences in some detail.

The small firms averaged 11.0 employees in the initial month, and 11.1 in the ending month. Therefore a change of one employee on average is a 9% change; and a one-person error in estimating either the current or future employment also is a 9% error. Thus for the small firms, small unit changes or errors are large percentage differences that have a major impact on the values for all of the relative, percent measures.

For the large establishments, the means for initial month, projected ending month, and actual ending month employment are 240.6, 239.3, and 236.7, respectively. Therefore a one-person change or error is only 0.4%. Stated differently, a 9% error for the average large firm is 22 employees. These substantial size differences make it inevitable that the percent errors will be larger for small firms than for large ones.

Another difference arises from the definition of no-change as an exact unit match between actual initial month employment and ending month employment. With an exact match, small firms' outcomes are no change 40.8% of the time; but for large firms, only 9.1% of the outcomes are no-change. In an earlier presentation of this analysis, we were criticized for requiring an exact match for large firms that predicted no-change. Therefore in this paper, a large-firm prediction of no-change is treated as a correct outcome, with zero error, if the actual percent change is less than plus or minus two percent. That adjusts the no-change outcomes for large firms upward by 7.7%, to a total of 16.8% – and also increases by 7.7% the percent of their forecasts with the correct sign.

Even with that difference in the treatment of no-change, a no-change forecast is more rational, and more likely to be accurate, for small firms than for large ones. Is higher accuracy due predominantly to a large number of accurate no-change forecasts evidence of superior forecasting ability? A naïve no-change forecasting methodology may yield the same degree of accuracy. Henri Theil examined this issue in detail in his analysis of the predictive value of anticipatory survey data, and his observation is particularly relevant here:

> *Consequently, if our variable happens to be characterized by a large percentage of no-change realizations, chances are that this raises the percentage of correct no-change predictions simply because no-change forecasts still are more frequent than no-change realizations are. This raises the proportion of correct forecasts; and this result is due, not to better forecasting, but to the observed distribution of change.... It is therefore conceivable that the superior performance for [the variable with a high percent of no-change outcomes] compared with [the variable with a low percent of no-change outcomes] has nothing to do with the "real" quality of the forecasts. (Theil, 1966, p. 365).*

In sum, the large differences in the percent of no-change outcomes and the potential impact of no-change forecasts for small versus large firms have a prominent bearing in assessing the forecasting ability of the two types of firms.

Key error measures for small and large establishments are given in Table 4. The MAPCE for small firms is 19.1%, versus 9.3% for large ones, but the mean absolute actual percent changes follow that same pattern, 18.1% and 8.3% respectively. Therefore the differences between the MAPCE and MAAPC are about the same 1%, and the Theil's U of 1.01 for small firms clearly is below the 1.14 for large ones. For us, these relationships are a good illustration of the value of using multiple statistics summarizing the percent change relationships, to obtain a better perspective about the forecasting accuracy.

The error measures for direction of change also require more than a cursory review. The 49.3% correct signs and Theil's Q of .46 for small firms in Table 4, versus 50.5% and .37 for large ones, appear to show that small firms' forecasts are as good as, or better than, those by large firms as signals of direction. However, examine the full Prediction-Realization results in Tables 5 and 6.

As shown in Table 5, the 49.3% correct for small firms come predominantly from 33.3% being correct no-change forecasts. Another way of stating their results is that they correctly predicted 33 out of every 41 no-change outcomes. However, they predicted no change 69.1% of the time, with over half of these predictions being wrong.

Regarding increases and decreases, they only correctly predicted 9 out of each 30 increases, and 7 of each 30 decreases. Therefore, what moderate forecasting success they have achieved stems nearly entirely from adopting a simple no-change forecasting approach. If the small firms cut in half their percent of no-change forecasts, redistributing them in proportion to their current pattern of increase and decrease predictions, their percent correct and Theil's Q would remain about the same.

The large-firm predictions of direction are a quite different situation. Their percent correct is a relatively low 50.5%. It is true that they also are predicting no-change excessively, 29.8% of the time when the realization is only 16.8%, as reported in Tables 4 and 6. However, this excess of no-change forecasts is very detrimental to their percent correct since they are wrong 64% of the time when they predict no change. When they do predict an increase or a decrease, the large firms are more accurate than the small ones. If they reduced their use of no-change predictions to zero, a proportional redistribution of their forecasts between increases and decreases could dramatically increase their accuracy. Such a redistribution would push their percent correct up from 50.5% to over 60% – considerably higher than that of small firms – and their Theil's Q from .37 to at least .52.

In sum, if small firms merely reduce their use of the no-change prediction, their overall accuracy, at least for direction of change, is not likely to improve. They also will need to develop a better forecasting process that increases their ability to forecast that an increase or decrease will occur. On the other hand, large firms can increase their accuracy in forecasting direction of change and level of employment merely by reducing the percentage of the time that they adopt a no-change forecast.

Table 5 - Small Establishments, 1 to 49 Employees, All Quarters
Sample Size 13,996

Error Measures for Three-Month Percent Change

Size of Percent Change	
Mean Predicted % Chg., MPC	3.3%
Mean Actual % Chg., MAC	3.9%
Mean Percent Change Error, MPCE	-0.5%
	0.0%
MAPCE	19.1%
Mean Abs. Actual % Chg., MAAPC	18.1%
Theil's U	1.008
Theil's Q	0.463
(0-1 Index for Quality of Change Prediction)	

Prediction-Realization Table

	Realization			
Prediction	Increase	No Change	Decrease	
Increase	9.1%	4.8%	4.8%	18.7%
No Change	17.8%	33.3%	18.0%	69.1%
Decrease	2.7%	2.7%	6.8%	12.2%
	29.6%	40.8%	29.6%	100.0%

Correct Sign	Increase	No Change	Decrease	
	9.1%	33.3%	6.8%	49.3%

Error by Over/Under	Correct %	Underest.	Overest.	
	36.7%	48.2%	15.1%	100.0%

Table 6 - Large Establishments, 50 or More Employees, All Quarters
Sample Size 1,851

Error Measures for Three-Month Percent Change

Size of Percent Change	
Mean Predicted % Chg., MPC	-0.2%
Mean Actual % Chg., MAC	-1.0%
Mean Percent Change Error, MPCE	0.8%
MAPCE	9.3%
Mean Abs. Actual % Chg., MAAPC	8.3%
Theil's U	1.142
Theil's Q	0.370
(0-1 Index for Quality of change prediction)	

Prediction-Realization Table

	Realization			
Prediction	Increase	No Change	Decrease	
Increase	23.6%	3.9%	16.4%	44.0%
No Change	9.0%	10.6%	10.2%	29.8%
Decrease	7.7%	2.3%	16.3%	26.3%
	40.4%	16.8%	42.8%	100.0%

Correct Sign	Increase	No Change	Decrease	
	23.6%	10.6%	16.3%	50.5%

Error by Over/Under	Correct %	Underest.	Overest.	
	12.5%	50.9%	36.6%	100.0%

Which group has the better forecasters? Based on the whole set of error measures, our judgment is the large establishments have an edge, but not a very large one at the present. Whatever one's decision, the nature and results of the Illinois forecasts definitely are different for small firms and large ones. So different that showing the distributions of past realizations separately for small firms and large ones in reports to establishment forecasters may change the forecasting strategy of each group in a way that increases the overall accuracy of the survey forecasts.

Conclusions

Predicting Total Employment

The 14 recent Illinois surveys of employers provide percent change predictions for total Illinois employment that are more accurate than no-change forecasts once the nine industry sector predictions are appropriately reweighted. The survey predictions do have an underestimation bias, both in size (MPCE of –0.5%) and number of underestimates (64.3% under versus 35.7% over). However, comparing the survey-based predicted growth rates to the naïve benchmark, the survey predictions have a lower mean absolute percent error, 1.3% versus 1.6%; and a Theil's U of .868.

The direction of change also is important information. The survey-based predictions for total employment had the correct sign 64% of the time (versus naive forecasts only 7%). The Theil's Q value of .59 indicates that the survey's past direction signals are relatively useful information, beyond just knowledge of the prior outcome distributions.

Significantly, the 36% of the time when the survey forecasted a gain, the survey was 100% correct. Therefore knowing that positive signal is a definite information gain. *Predictions of declines in total* employment were not reliable, only correct 44% of the time, so survey forecasts of decreases must be weighed against other sources of information. However, the analysis by sector shows the low accuracy in predicting declines stems predominantly from one sector and may be reduced by reviewing the sampling and verification procedures.

Our judgment, based on this evidence, is the employer surveys have value as a methodology for forecasting short-term percentage and direction of change in total state employment.

Predicting Individual Establishment Employment

The stable nature of the survey respondents' employment over the three-month horizons with seemingly random errors is an environment where most paired differences tests showed no significant differences exist. This environment causes the evaluation of the firms' forecasting accuracy to be more subjective than had been hoped, but also increases the value of utilizing a number of error measures.

Error measures for the full set of 15,847 establishment forecasts show that the individual forecasts on average are no better than no-change forecasts in predicting the size of the firm's future employment and the percent change from current employment.

As direction signals, only 48% of their forecasts had the correct sign. However, the prediction-*realization analysis shows establishments' over-reliance* on no-change as their forecasting strategy is a major weakness. Providing them with domain knowledge about the distributions of past outcomes may lead them to adjust their forecasting strategies in a way that increases the individual establishment's accuracy.

The analyses of subsets of the establishments by season of the year and by seasonal industries versus nonseasonal ones revealed no significant differences in accuracy attributable to seasonality, other than the March-to-June errors being somewhat lower than the other three periods. Moreover, the evidence suggests that establishments in highly seasonal industries are not effectively recognizing and incorporating their seasonal variations into their employment forecasts.

Analysis of small versus large establishments indicates that the large establishments have a slight edge in forecasting ability. However, their past errors have been larger than naïve forecast errors. Moreover, the large establishments' signals of direction of change have been only marginally better than those of small establishments – hampered also by excessive reliance on no-change forecasts.

In sum, the establishment forecasts and outcomes imply the average Illinois establishment is *not using effective short-term forecasting procedures.* Enhancing their forecasting skills could be beneficial to the establishment as well as improve the accuracy of the employer surveys.

REFERENCES

Fildes, R. and S. Makridakis, 1995, "The Impact of Empirical Accuracy Studies on Time Series Analysis and Forecasting," *International Statistical Review,* 63, 3, 289-308.

Hibon, M., and S. Makridakis, 1999, "M3-Competition", unpublished paper presentation at ISF 19.

Theil, H., 1966, *Applied Economic Forecasting,* Amsterdam, The Netherlands: North-Holland.

DATA OBSOLESCENCE AND FORECASTING

Othmar W. Winkler
Professor emeritus
Georgetown University
FFC/Sept. 14, 2000
winklero@georgetown.edu
winklero@msb.edu

The Forecaster and the Past

I will not concern myself with methods of forecasting but with the statistical data-inputs on which such statistical forecasts are based[1]. I will be concerned with the potential for prediction, the predictive value, of socio-economic data. When the results of forecasting models are evaluated, their lack of success often is blamed on the 'lack of precision of the data,' or as 'the influence of a human factor in forecasting.[2]

The term 'predictive value' is sometimes understood to mean the ability of one time series to give advance notice of changes in another series which lags behind[3]. I am using 'predictive value' of data broadly as the potential to anticipate future socio-economic developments.

Every statistical figure dealing with society supposedly reflects the state of a socio-economic situation at a certain point in time. As that situation unfolds, the analyst keeps abreast of the changes with new data. When the availability of a statistical figure is delayed, the situation it describes corresponds to the actual state of affairs at the moment of analysis only to the extent to which that situation has not changed. Unfortunately, in this fast developing society something of the relevance, for describing the current situation, even of the most recent data, has already been lost by the time those data become available.

Imagine, for example, how useful the time series of the index of industrial production would be to a forecaster who at years' end receives the August figure as the latest available datum. How well would s/he be informed about the production situation at the end of that year? How useful would that time series be in a forecast for the first or even later semesters of the following year? Or how useful can the information contained in the 1990 population census be to a forecaster who must rely on it as the only available information in 1996, years after that census had been taken? Evidently there is a point in time beyond which a statistical figure ceases to be of value in

---

[1] I have presented an earlier version of the following ideas as "Forecasting: the predictive Value of Statistical Data" in Proceedings of the Business and Economic Statistics Section of ASA Washington, D.C. 1968, pp.381-385. Also as "The Effect of Data Obsolescence on Economic Forecasting - A special Case of Timeliness" in: Contributed Papers, ISI, 46th Session, Tokyo 1987, pp.473-4

[2] e.g. James B. Wong, ~~Business Trends~~ and Forecasting, an annotated guide to theoretical and technical publications and to sources of data, Gale Research Co. Detroit, Mich, 1966, p.31 and Walter E. Hoadley Jr. "The Importance and problems of Business Forecasting" in: Herbert Prochnow, ed. Determining the Business Outlook, Harper Brothers, New York, 1954, p.23

[3] See e.g. Milton H. Spencer, Colin G. Clark, Peter W. Hoguet, Business and Economic Forecasting, pp. 202, 203.

assessing the current, let alone the future state of a socio-economic situation. This important fact of "data obsolescence" is of much less importance in the data -mostly accurate measurements - in the physical/natural sciences. It is because of statistical theory's heavy reliance on the methods developed in those hard sciences that economists and social scientists have not paid much attention to the fact of data obsolescence and its consequences. The assumption of continuity in patterns and relationships which underlies every forecasting method must be understood in the light of this basic fact[4].

A good part of forecasting consists in understanding the past, tracing down the historic roots of the social and economic forces that are responsible for the present state of the situation in order to extend these into the future. Such an understanding of the history of a situation lies at the heart of the matter. The forecaster must learn to understand that the different parts of his historic data in a time series are of different value to him: evidently he should pay more attention to the newer, more recent figures, than to the older ones. Usually he should confine his attention to a rather limited time-span. As time moves on, that time-span also keeps advancing. S/He must not conceive of a socio-economic time series as an ordinary climbing vine that continues to grow at the tip of its runners while remaining fully alive in all its parts. He must instead conceive of it as one of those rare creepers, the older parts of which die off gradually while it continues to sprout new leaves and roots at the tip of its runners, clinging to the new ground and feeding on it. The forecaster must not burden his/her work with data that have become obsolete, and therefore irrelevant for anticipating the future developments of the socio-economic situation to be forecast.

Statistical obsolescence, its causes and assessment

Forecasters have long recognized the need for rapidly available figures and were willing to trade off loss in accuracy against timeliness. The custom of e.g. the BLS to present their published price and productivity data in such a way that the newest figures are listed first, then the older ones, in reversed time sequence, and limited to relatively short time spans, therefore makes good sense.

The awareness has not yet sunk in that socio-economic data over time become useless. They expire, so to speak, like dairy products or medicines. regardless of their original cost. The process of obsolescence in the data, the fading-out of descriptive value through the loss of timeliness, continues at an uneven speed. No fixed formula can do justice to this loss. After some time every statistical figure has become valueless for understanding the present situation, let alone its future. All expired socio-economic data have become useless and are to be discarded from the forecasting process.

Statistical obsolescence stems from changes in the underlying causal system, and is due to factors that are internal and external to the socio-ecnomic situation to be forecast.

---

[4](The theologian Paul)" Tillich maintains that humans were never able to bear the thought of having their experience thrust into a past where it would be totally lost. And this is the reason why they have always sought ... to erect obstacles to the diminishment of their memory (p.114) ..it is extremely difficult to imagine how anything could be imbued with lasting significance..(p115)" John F.Haught, The Cosmic Adventure-Science, Religion and the Quest for Purpose, Paulist Press, New York, NY 1984

Internal factors act whenever individual workers, business firms, fixed capital assets, etc. are being replaced by new, different ones which can perform at higher levels of quality and quantity. These internal factors cause changes, through the superior preparation of the new entrants in the labor force, installation of computers and robot systems that increase productivity and the superior efficiency of new ways of managing business organizations. The effects of such innovative changes often are not directly reflected in the data on production, exports, etc. Besides the outright replacement of the workforce and of equipment, there are many small changes. Older workers are retrained, new concepts of depreciation are introduced, business transactions are made faster and cheaper, 'flextime' and other new management strategies alter the responses of the socio-economic actors to the customary incentives of society. Such ubiquitous 'internal changes' cause a creeping loss of continuity in all data dealing with aspects of society.

External factors are those that refer to broader changes in the general socio-economic setting, the change from war to peace production, racial integration ordered by law, Title IV legislation dealing with sex discrimination in employment, changes in the interest rate by the Federal Reserve Board, and every change in existing government regulations that affects the industry or region for which a forecast is to be made.

Every indication of changes, then, is also an indication of additional statistical obsolescence in the data that were obtained before that change. These shifts in the combination of socio-economic forces are gradual, seldom noticed in the data. Only few changes in the social and economic environment leave visible marks in the data. Obsolescence works as an unspectacular erosion that will not become visible in the figures. This unpredictable process of becoming irrelevant takes place with uneven speed, constantly changing within the same series. It is here that prudent judgement of the perceptive forecaster must enter. Obsolescence is at work in all statistical data, affecting the relationship between time series, and aggravating the problems of 'proxy series' and of those series that are difficult to interpret because of methodologic changes in data gathering methods or changed definitions[5].

Despite numerous hints to the great need for staff and upkeep of the forecasting models in the description of actual forecasts, statistical obsolescence is hardly ever explicitly considered.[6] Although forecasters may have been aware of its presence little seems to have been done about it.

Obsolescence in data leads to the important question: How far back can data be used as inputs into a forecasting model? Obviously there is no pat answer available. The forecaster will have to study each situation to be forecast. All events in society that may have affected the continuity of the causal system which underlies the socio-economic situation to be forecast may have to be investigated and judged for its impact on the data at hand. That task pertains to the economist, engineer, manager, sociologist, demographer, in short, the expert in the subject

[5] Spencer, op. cit. p. 91.

[6] Spencer, op. cit., p. 20, 21 and 35.

matter, not the statistician! That expert must appraise the importance of the loss in continuity, *assessing as objectively as possible, how much of the continuity of each figure has been lost* during each time period for which data are provided. She/He will determine to what extent the data are still relevant. For forecasts performed continuously, the relevance of each figure will have to be reassessed and the assigned weights be adjusted for each new forecast to be made. It is important that this is done by informed expert judgement, not mechanically by a fixed (mathematical) formula. Obsolescence of data is to be estimated as the amount of 'lack of continuity' in the underlying technical, social and economic conditions that connect the situation in the earlier period to the present.

Such an assessment of gradual discontinuity should be indicated by weights. These weights *assessing the loss of continuity in the underlying socio-economic situation, could be expressed by* decimal fractions, appearing like probabilities. These weights, relating the degree of obsolescence, will lie between 1 and 0. A weight of 1.0 would indicate that no changes in the internal and external factors could be found between two time period under investigation. If a socio-economic situation has changed completely, the factor expressing continuity would become zero. Such a weight would then also be assigned to all data in a time series before the one with 0.0 continuity..
An estimate of a joint continuity factor of say, .10 - that would be an obsolescence factor of .9 - simply means that the information gleaned from the figure of that time period should be used in forecasting with only 1/10 of the importance given to the figure from the present time period.

As an hypothetical example, assume a time series that goes back to 1979. Assume also that a competent staff of analysts has studied closely that series. These subject-matter experts assess the degree of loss of continuity of the figures of that series, for each year, relative to the previous year. To express the degree of continuity, or the lack of it - loss through obsolescence - each expert assesses that continuity as a decimal fraction between 0 and 1. After discussion, we assume that these experts have agreed on the continuity ratings, given below. These continuity ratings are not to be mistaken for probabilities. A rating of 1 would signify that the subject-matter experts - not the statistician - found no indication that the conditions in that industry have changed. A rating of 0 would indicate a complete rupture in the conditions between two consecutive time periods. The figure of that period, and all earlier data of such a time series, would have been found to be useless for forecasting. Suppose that the degree of continuity -- or the lack of it as obsolescence – was determined between each two subsequent yearly data as follows:

```
Year 1979 '80 '81 '82 '83 '84 '85 '86 '87 '88 '89 '90 '91 '92 '93 '94 '95
          \ / \ / \ / \ / \ / \ / \ / \ / \ / \ / \ / \ / \ / \ / \ / \ /
Continuity .95 .98 .80 .75 .79 .40 .35 .60 .80 .86 .91 .97 .99 1.0 .90 .94
```

The figure ".95" for 1979/80 would indicate that between 1979 and 1980 the situation underwent only minor changes. The high stability in the situation was assessed as .95, a stability of 95% with a loss of continuity of about 5%. Between the years 1984/85, in contrast, major discontinuities in that industry were observed, leaving only 40% of the conditions to carry over into the following year. This low continuity corresponds to a loss through 'obsolescence' of about 60%. On the other hand , there were no changes observed between 1992 and 1993. The continuity ratings of that series will be shown, in reverse order, to determine the joint "discounts" for obsolescence, beginning with the most recent figure:

Year	Determining Joint Continuity Ratings for 1996-97		Obsolescence
1995 -96	1.0	= 1.00	.000
1994-95	1.0*.94.	=.940	.060
1993-94	1.0*.94*.90	=.846	.154
1992-93	1.0*.94*.90*.1.0	=.846	.154
1991-92	1.*.94*.90*1.0*.99	=.838	.162
1990-91	1.0*.94*.90*1.0*.99*.97	=.812	.188
1989-90	1.0*.94*.90*1.0*.99*.97*.91	=.739	.261
1988-89	1.0*.94*.90*1.0*.99*.97*.91*.86	=.636	.364
1987-88	1.0*.94*.90*1.0*.99*.97*.91*.86*.80	=.509	.491
1986-87	1.0*.94*.90*1.0*.99*.97*.91*.86*.80*.60	=.305	.695
1985-86	1.0*.94*.90*1.0*.99*.97*.91*.86*.80*.60*.35	=.107	.893
1984-85	1.0*.94*.90*1.0*.99*.97*.91*.86*.80*.60*.35*.40	=.043	.957
1983-84	1.0*.94*.90*1.0*.99*.97*.91*.86*.80*.60*.35*.40 *.79	=.034	.966
1982-83	1.0*.94*.90*1.0*.99*.97*.91*.86*.80*.60*.35*.40 *.79*.75	=.025	.975
1981-82	1.0*.94*.90*1.0*.99*.97*.91*.86*.80*.60*.35*.40 *.79*.75*.80 . . .	=.020	.980
1980-81	1.0*.94*.90*1.0*.99*.97*.91*.86*.80*.60*.35*.40 *.79*.75*.80*.98	=.020	.980
1979-80	1.0*.94*.90*1.0*.99*.97*.91*.86*.80*.60*.35*.40 *.79*.75*.80*.98*.95	=.019	.981

These ratings indicate the cumulative effects of obsolescence of the earlier data for any attempts to anticipate in 1995 the scenario of the socio-economic setting for this series in 1996 and beyond. These figures indicate that in 1995 the 1990 figures of that series can be relied only with 73.9% of their value when trying to forecast beyond 1995. That 73.9% implies a 100% - 73.9% = 26.1% loss of continuity due to data obsolescence, informing the forecaster that these older data are not to be used at a par with the latest figures, but with the indicated amount of "discount for obsolescence." The 1987 figures should be used for forecasting purposes with only 30.5% of their original value.

Although the more knowledgeable, perceptive and gifted forecaster will produce better forecasts, the final determination should be achieved by discussion and consensus between the members of a team of subject-mater experts charged with assigning weights for obsolescence to the data.

Changes in the underlying causal system have been measured before[7]. For purposes of forecasting, however, a more sensitive perception of changes, and of their impact, is required. It

[7]See e.g. Gregory C. Chow "Tests of Equality between Sets of Coefficients in two Linear Regressions" Econometrica, Vol. 28, 3, July 1960,
Also: "Das Lexis'sche Dispersionsverfahren," in: Wilhelm Winkler Grundriss der Statistik I, Wien 1947, Manzsche Verlagsbuchhandlung, p.73-79.

is obvious that data before the Korean war in 1952 should not be used any longer for forecasting. This goes against the widely held, mistaken belief among forecasters, that longer time series give better forecasts because supposedly "there is strength in numbers." True, according to sampling theory, larger samples allow more reliable conclusions about a population. But the data of social and economic time series are not drawn at random from a timeless population. Moreover, each figure usually is itself a population describing the successive stages in the development of a situation in society. Without being aware of it, I believe that statisticians' thinking today is still dominated by the concepts of statistical sampling and inference. All statistical data, time series included , are treated as if they were random samples. Yet the data of most socio-economic time series are not a set of simultaneously existing sample units. Instead we must realize the true, descriptive nature of socio-economic statistical data, which, rightly understood, leads to limiting their use only to the relevant, more recent data. In other words, using only relatively short portions of the available time series. The intuitive understanding of this fact seems to account for the popularity of exponential smoothing in forecasting.

Obsolescence and size of the aggregate

A question that has been raised repeatedly: can the forecast of a time series be improved by combining the forecasts of its sub- time series? Is such a combined forecast superior to a direct forecast of the larger aggregates in a time series?

Time series consisting of large aggregates describe a less pinpointed, broader picture. Such series show only those major net-changes in the socio- economic situation that reach beyond the aggregation limits with regard to time interval, subject matter and geographic territory. Everything else in these large aggregates has been eliminated by internal compensation. As a result time series of large aggregates fluctuate less, nor do these become obsolete as rapidly as the data of small aggregates. The broad picture, that large aggregates describe, is less affected by the innumerable day-to-day changes that occur in small regions and narrowly defined subject categories.

These same day-to-day local changes do affect time series of narrowly defined aggregates. They fluctuate more frequently and more strongly, reflecting the minor changes in the business scene with greater immediacy. Consequently they become more rapidly obsolete and their forecasting span into the future is much shorter. Because this does not allow to trace the present situation very far into the past, their forecasting range is correspondingly short, allowing only short-term forecasts. Time series of wider aggregates -- wide with regard to their geographic territory and/or the length of the time period and/or the width of definition of the subject-matter -- have a lower rate of obsolescence and permit longer-range forecasts than those based on more narrow aggregates. If various such short-term forecasts are combined into a forecast of the total series, such a combined forecasting range does not extend farther into the future, as the forecasting range of the other component series. It will not allow forecasts as far into the future as the forecasting span of the time series formed by aggregation of the smaller component time series. In light of these facts, combining the forecasts of the part-series of an aggregate will not improve the longer-range forecasts made with the aggregate series[8].

When the relationship between various time series is explored with n-dimensional multivariate analysis, the rates of obsolescence of these n time series may differ. In that case the joint obsolescence for the data of a given time period is the product between the individual obsolescence ratings determined for each series for that particular time period. These obsolescence factors are to be used like frequency-weights with which each of the multidimensional points on the regression surface are to be weighted, when calculating regression parameters. This measure of obsolescence is an attempt at quantifying the impact of historic developments on the usefulness of older data.

Some conclusions

Few forecasting models have consistently performed well. The reasons, I suspect, were not necessarily the faults in the economic logic on which they rest, but the indiscriminate input of data. All models will improve their performance if their parameters are computed with proper regard for statistical obsolescence.

When adjusting seasonal fluctuations by electronic computers, earlier models had limited data storage capacity, in many cases capable of accommodating time spans of not more than 15 years. This was deplored as a drawback[9]. In reality, such a limitation really may have been a blessing. A span of 15 years is probably more than is needed for most forecasting purposes in these times of rapidly changing technology.

This discussion may also have practical consequences for the storage capacity of data banks. Obsolescence should lead to a frequent turnover within the storage area of the bank. As soon as data begin to expire beyond the point of high usefulness, they ought to be transferred from the more costly 'interactive storage area' into cheaper, less readily accessible storage areas, and finally, into 'dead-data files.' Such frequent, obsolescence-based rotation should alleviate storage problems and lead to a more economical use of electronic data storage[10]. Compromises though, will have to be made between uses of data whose component series have different expiration ranges.

Another conclusion is of a more academic nature. When e.g. for the purpose of determining insurance rates, relative frequency distributions are computed from time series. Statisticians leaning toward the "objective" interpretations of probabilities would include as many data of the time series as possible. In a situation of rapid change and obsolescence, these 'probabilities' may be based on a fairly short part of the time series, approaching in the limit, "subjective"

---

[8] See e.g. David C. Melnikoff, "Long Term Projections and Business Decisions" ~~Proceedings~~ of the American Statistical Association, 1957, Business and Economic Statistics Section, p. 337 upper right.

[9] Julius Shiskin, Harry Eisenpress "Seasonal Adjustments by Electronic Computer Methods" NBER, Technical paper No. 12, New York, National Bureau of Economic Research, 1958 p. 427, especially his reference to Method I. The fact that the capacity of computers since then has been extended to 50 and more years does not change my point.

[10] Georgetown University Library has begun in 1999 to remove books for which there was only minimal demand, from its 'active' shelves at the library and store them in a geographically remote, less rapidly accessible, cheaper storage location.

probabilities. These can be understood as "probability distributions determine from time series with extreme obsolescence, thereby bridging the opposing views of subjective and objective probability.

But there is also another practical side to this: Insurance companies determine the ratios of insurable events from time series - ratios incorrectly referred to as probabilities . These are only rarely changed to adjust for major changes in society. But in fact all insurance rates should be re-calculated on a regular basis, from up-to-date, revolving sets of data, that include the newest data while gradually eliminating obsolete data that no longer represent the social, demographic and economic reality, e.g. for life insurance purposes. That would be another important application of the proposed adjustments for data-obsolescence.

Although forecasting is a necessity, nobody really can predict the future. We were reminded of this by the world oil crisis of 1973 that caught the world by complete surprise. That is bound to happen again because forecasting with statistical data is like a person who advances with his/her back to the direction in which s/he intends to move. Instead of looking forward, watching where s/he is going to step next, the forecaster looks back, searching for clues to the future in the past, relying on the statistical records of the past for hints about future developments[11].

Despite such pessimism, a plausible defense, for the frequent case of forecasts that missed the mark, could be as follows: Assuming you developed a perfect forecasting model that gives unfailing results. Two things are bound to happen. 1. As forecasting is not a spectator sport, but made to guide action, those who ordered the forecast will take advantage of that predicted boon or act to ward off the predicted threat. And 2. Other forecasters also will have made forecasts. Even if those were not as good as your's pro-active action will be taken based on their forecasts. By the time the predicted future arrives, it has been tampered with to such an extent, thanks to all these forecasts, that it became something quite different from the future that had existed at the time when you made your forecast. So you can feel vindicated about the quality of your own forecast: it would have been perfect if everybody just had left that future alone!

[11]Peter Drucker summarized this succinctly: "We must start out with the premise that forecasting is not a respectable human activity and not worthwhile beyond the shortest periods. Strategic planning is necessary precisely because we cannot forecast" Management, Tasks, Responsibilities, Practices, Harper & Row, New York, 1974, p.124

Concurrent Sessions II

ISSUES IN EMPLOYMENT PROJECTIONS

Chair: Norman C. Saunders
Bureau of Labor Statistics, U.S. Department of Labor

The Impact of Changes in Both Final and Intermediate Demand on the Structure of Industry Employment, 1978 to 1995,
Art Andreassen, Bureau of Labor Statistics, U.S. Department of Labor

Business Inventory Practices: Model and Analysis,
Jay Berman, Bureau of Labor Statistics, U.S. Department of Labor

Modeling the Demand for Skills,
Charles Bowman, Bureau of Labor Statistics, U.S. Department of Labor

The Impact of Changes in Both Final and Intermediate Demand on the Structure of Industry Employment, 1978 to 1995

Art Andreassen
Office of Employment Projections, Bureau of Labor Statistics

The distribution of employment by industry has undergone major shifts since 1978 and number of explanations have been advanced as to why this has occurred. The following study investigates and measures the extent to which changes in demand have contributed to this new employment distribution. Between 1978 and 1995, a growing and fully employed labor force together with increases in productivity drove output higher and added 31.9 million new jobs to the 98.1 million that were already in existence. This healthy 32% growth in employment was not shared equally by all industries. Many industries in fact had employment declines as changes in products, production processes, management practices and tastes redistributed the job structure. Changes over time in the share taken by an industry's employment is due to the interaction of productivity and demand. Productivity impacts employment levels by affecting an industry's relative price and its use of labor. Relatively higher productivity growth allows an industry the option of raising prices less than other industries increasing the demand for its output. Alternatively, a relatively high rate of productivity growth permits an industry to increase output with a lesser increase in employment. Manufacturing industries have high rates of productivity growth which allow output increases faster the employment increases. This study however concentrates on the other factor, demand, and the contribution made by each of its two components, demand by final users and by intermediate users.

Employment by Industry

The economy can be divided into a goods producers, composed of agriculture, mining, construction and manufacturing, and service producers composed of trade, transportation, communications, public utilities, services and the government. Year after year, in good times and bad, a common refrain has been that jobs producing goods are disappearing while those producing services are increasing. And in truth, whether the economy is at a cyclical peak or trough, the manufacturing sector is losing a share of employment to the advance of the service sector. In both peak years (1979, 1989) and trough years (1982, 1992) manufacturing jobs have represented a declining share while service jobs have continuously trended in the opposite direction.

Moving from the composite sectors to the subordinate industries within reinforces the extent to which service industries are growing and manufacturing industries are declining. Of 97 manufacturing industries , 67 had job declines while only 2 of 38 service industries fell. Thirteen service industries had an increase in the number of jobs twofold or more with the personal supply industry increasing over fivefold. The agriculture sector's positive growth is due to the more than doubling of the agricultural services industry, an increase which swamps the much larger agriculture production decline. This is a common theme running through the comparison of these two periods: movement from the production of things to the provision of services. Following up on this concept, all of the mining industries suffered losses and construction, although displaying growth, did so at a slower rate than the total. The classification scheme we are using harks back to a time of smokestack production, since it over-represents manufacturing and under-represents services. In durable manufacturing only 11 of 38 showed any growth with 3 growing faster than the total. The medical equipment industry is the only manufacturing standout with 51% growth. Nondurables fared slightly better with positive growth in 20 of 39 industries of which only 5 grew faster than the total.

Employment in the railroad and water transportation industries dropped by more than one half reflecting the increased productivity of containerized shipping. On the other hand the air transportation, the passenger transportation and the miscellaneous passenger services industries more than doubled as deregulation led to higher productivity, lower prices and increased demand. Wholesale and retail trade just about as fast as total while the eating and drinking places industry had an increase more than twice the total. In services the personal supply services industry- temporary help- had the largest growth in the economy, a 5.6 fold increase, and in numbers was only surpassed by the much larger trade industries. The computer and data processing service industries was second in growth rate with a five fold increase.

Sources of demand

Employees produce output to fulfill the demand of either final users or intermediate users. Intermediate demand is purchases by other industries to be used as inputs for further processing. Changes in taste and in technological processes alter the demand distribution over time. Both final and intermediate demand respond to evolving economic conditions such as new sources of supply, different relative price movements and advances in knowledge. Viewing employment distributions at the terminal points of a period exposes the combined extent of all changes but offers little insight into the distinct impact of each type of demand. An attempt is made here to parse and measure the unique contribution each of the two types of demand.

Variations in the production process and thus to intermediate demand are much more gradual and, while not entirely immune, are less sensitive to the immediate economic climate. The production process responds especially to two influences, new products and new management practices. Each could not be studied individually because it was difficult to measure their specific contributions. Further, both induce similar responses on the part of business enterprises. With the introduction into the production process of a new product there is an increase in an input used by the purchasing industry's employees. This new input would not be introduced unless it's cost were more than compensated for by savings in other areas of production. Employment in the purchasing industry will decrease if the new input increases productivity but not output, employment will not be affected if the change in both output and productivity are equal, or employment will increase if output increase more than productivity. Employment and output in the supplying industry will increase in response to increased sales.

On the other hand, new management processes, such as those which have been occurring under the rubric of re-engineering, have more pervasive impacts on the production process. Rather than just purchasing a new product and incorporating it into the existing input structure an enterprise will purchase a procedure that replaces a large block of inputs, labor included. Even if the output of the purchasing industry remains unchanged the number of employees will decline since the new procedure includes labor now located in the supplying industry. When a manufacturing industry replaces a function usually done in-house by inputs from outside suppliers, employment will move from the purchasing to the supplying industry. The supplying industry may use the same material inputs but it combines them with its own employees. Within the purchasing industry there is now one material input purchase, from the supplying industry, which replaces the previous separate inputs, including the labor. Although this is a development that has occurred in the past it has taken place mainly among manufacturing industries and employment has shifted among manufacturing industries. Recent changes in computers, telecommunications and transportation are different such that manufacturing industries now purchase more input replacement from outside the manufacturing sector, especially from services. The input/output system can be used to measure the amount of the shift of employment to other sectors due to the combined action of outsourcing and re-engineering.

Calculation of the Impact of Demand

As part of its projection process the Bureau has created an historical database which has been utilized to test these different scenarios on employment. Using this database with assumptions about what has happened over the past one can measure explanatory theories and test their veracity and their relevance. This calculation can be carried out because the Bureau's projection's process entails distinct steps and each can be varied separately to test different scenarios. This study relies on the historical data series of final demand and of input-output tables. An input-output table measures the material inputs purchased from all the other industries within the economy as well as the factors of production that are necessary to produce the purchasing industry's output. This system measures not only the first level of purchases but also the production induced in other industries to produce this first level of purchases. Thus, final demand impacts industry employment not only by buying directly from that industry but also by inducing output in those industries which provide inputs into the production process of that industry. Supplementing the input-output tables are measures of industry productivity that allow conversion from industry output in dollars to industry employment. An input/output table from one year can be combined with the different final demands to compare the industry production necessary to produce it. It is this capability to combine different production processes to different demand structures from which to glean insights into the evolution of the present employment distribution. Actual 1995 industry employment is compared, first with the employment necessary to satisfy a 1978 distribution of demand and then with the employment a 1978 technology would have generated to satisfy the demand of 1995. The differences in the number of jobs generated by each of the two scenarios when compared to actual indicates the relevance of the effects of changing taste and changing technology.

Generated employment at the sector level shows that it is shifts in both final and intermediate demand acting together that has fashioned the 1995 employment structure. Presented are comparisons of actual employment with that resulting first from 1978 final demand then from the older production process and finally the combined result. Both types of demand individually contribute fewer manufacturing jobs and more services. This study does not attempt to explain the causes of the 32 million increase in jobs but does try to quantify the effect of the evolution of taste and technology. A 1978 final demand distribution satisfied by a 1995 production process would create over 588 thousand more jobs since government would have had a 3.5 million more jobs while the private sector would have had almost 3 million less. On the other hand, using the 1978 production process to satisfy the 1995 final demand would generate almost 1.5 million less jobs but manufacturing and construction would require 2.2 million more. Together the 1978 demand distribution and production process would have generate 845 thousand fewer jobs in 1995 since there would have been 4.6 million fewer in the private sphere with the 3.8 million more in the government. Although a 1978 final demand distribution or production process alone would result in more manufacturing employment neither alone would produce manufacturing jobs matching jobs the number actually obtained in 1978 emphasizing that it is not industry restructuring alone that is responsible for the shift to service employment. In fact, final demand changes account for half the shift.

At the industry level one gets the full flavor of the impact of each demand shifts. Government employment, both Federal and state and local levels, is solely a function of final demand and displays the second highest relative difference after construction. Not entirely surprising since the downturn in the 1995 share of GDP represented by defense spending did not start until a decade after 1978. Non-defense employment has also shown a slight actual drop also as government has a constant employment level as Gross Domestic Product has grown 43%. State and local government education employment is at an actual lower level reflecting the stability of the number of pupils as opposed to the growth necessitated to educate the baby boomlet generation. Construction has more employment under both of the older demand sources in that in 1978 demand for new construction was healthy and would remain so into the 1980's. Unfortunately it would prove to be so healthy that in 1995 the economy would still be trying to work off the excess floor space. Along with this exuberant building was a drop in the need for office space as more work was capable of being done off site thanks to improvements in telecommunications and computers. The 1978 production process calls forth more construction employment too but this from the maintenance repair industry. This is a good example of the impact of outsourcing, companies no longer perform this work in house with their own employees but contract it out to the building services and the agricultural services (landscaping) industries. These industries are credited with the employment as opposed to the companies that actually purchase and benefit from these services.

Under the old demand patterns very few manufacturing industries would have less employment, in fact, both patterns would have contributed equally to an increase in manufacturing jobs and as well as equally generating fewer jobs in services. Manufacturing industries which expectedly bucked the trend and increased their employment in today's economy are computer and office equipment, communications equipment, electronic components and accessories and medical equipment, instruments and supplies. From the opposite side employment in services would be less under the earlier demand structure. Stand out industries are computer and data processing services, 83% fewer, management and public relations, 73%, personal supply services, 62%, and research and testing services, 56%. Obviously, all these industries have been affected by the technological revolutions in computer and telecommunications and the new management practices that have swept the economy. Increased employment in the health and residential care industries, on the other hand, resulted from more consumer demand as the population has aged.

Conclusion

An attempt has been made to separate and quantify trends in an economy that has seen both a relative and numeric increase in service employment while the goods producing industries, manufacturing in particular have declined. Both final users and industrial users have responded to new products, tastes and business practices by substituting newer for older goods and therefore impacting the use of labor across the market. Most industries respond almost exclusively to either final or intermediate demand but the net result of the interaction of both demands on the economy has been the creation of fewer jobs in manufacturing and more in services. Much discussion has centered on the use by manufacturing of outsourcing and subcontracting as the source of the decline in manufacturing jobs. This study has demonstrated that almost half of the shift to service jobs since1978 has resulted from the changing pattern of purchases by consumers.

Employment by Industry, 1995

Actual,Recalculated and the Differences

(Thousands of Jobs)

	Actual	1978 Final Demand Distribution	Difference	1978 production process	Difference	Total Difference
Total	129,998.8	130,586.8	588.0	128,526.3	-1,472.5	-884.
Agricultural production......................	2,341.0	2,561.2	220.2	2,399.8	58.8	279.
Agricultural services........................	1,217.1	1,098.9	-118.2	711.2	-505.9	-624.
Forestry, fishing, hunting, & trapping.......	92.0	86.7	-5.3	94.0	2.0	-3.
Metal mining................................	52.3	35.6	-16.7	89.9	37.6	20.
Coal mining.................................	105.4	96.5	-8.9	108.4	3.0	-5.
Crude petroleum, natural gas, and gas liquids	161.9	201.7	39.8	242.8	80.9	120.
Oil and gas field services..................	169.3	413.3	244.0	201.2	31.9	275.
Nonmetallic minerals, except fuels...........	108.3	124.9	16.6	132.7	24.4	41.
Construction................................	6,632.6	7,836.2	1,203.6	7,607.0	974.4	2,178.
Logging....................................	132.1	156.2	24.1	153.2	21.1	45.
Sawmills and planing mills..................	194.7	214.1	19.4	207.4	12.7	32.
Millwork, plywood, and structural members...	291.8	325.3	33.5	298.8	7.0	40.
Wood containers and misc. wood products.....	160.2	166.2	6.0	153.9	-6.3	-0.
Wood buildings and mobile homes............	82.4	94.1	11.7	114.2	31.8	43.
Household furniture........................	293.8	357.3	63.5	349.7	55.9	119.
Partitions and fixtures....................	90.2	96.3	6.1	86.6	-3.6	2.
Office and misc. furniture and fixtures.....	149.9	118.7	-31.2	136.4	-13.5	-44.
Glass and glass products...................	154.3	165.9	11.6	204.0	49.7	61.
Hydraulic cement...........................	17.9	22.6	4.7	24.5	6.6	11.
Stone, clay, and misc. mineral products.....	177.2	188.6	11.4	262.9	85.7	97.
Concrete, gypsum, & plaster products........	208.2	241.8	33.6	230.9	22.7	56.
Blast furnaces and basic steel products.....	242.6	295.7	53.1	391.8	149.2	202.
Iron and steel foundries...................	131.1	171.0	39.9	234.1	103.0	142.
Primary nonferrous smelting & refining......	42.0	33.4	-8.6	84.5	42.5	33.
All other primary metals...................	44.7	39.8	-4.9	44.7	0.0	-4.
Nonferrous rolling and drawing.............	167.6	183.3	15.7	242.9	75.3	91.
Nonferrous foundries.......................	87.1	95.7	8.6	95.0	7.9	16.
Metal cans and shipping containers..........	41.3	42.8	1.5	65.9	24.6	26.
Cutlery, hand tools, and hardware...........	129.7	153.4	23.7	163.1	33.4	57.
Plumbing and nonelectric heating equipment..	58.0	74.2	16.2	75.9	17.9	34.
Fabricated structural metal products........	438.9	566.5	127.6	443.9	5.0	132.
Screw machine products, bolts, rivets, etc..	100.0	117.2	17.2	105.9	5.9	23.
Metal forgings and stampings................	252.7	282.7	30.0	286.5	33.8	63.
Metal coating, engraving,and allied services	131.2	128.5	-2.7	109.5	-21.7	-24.
Ordnance and ammunition....................	51.7	60.2	8.5	59.0	7.3	15.
Miscellaneous fabricated metal products.....	255.7	292.8	37.1	265.3	9.6	46.
Engines and turbines.......................	88.3	101.6	13.3	129.0	40.7	54.
Farm and garden machinery and equipment.....	101.2	202.7	101.5	75.4	-25.8	75.
Construction and related machinery..........	226.5	422.3	195.8	285.1	58.6	254.
Metalworking machinery and equipment........	351.5	498.7	147.2	367.0	15.5	162.

Special industry machinery..................	175.7	152.9	-22.8	195.7	20.0	-2.
General industrial machinery and equipment..	257.1	382.9	125.8	244.7	-12.4	113.
Computer and office equipment...............	353.2	93.6	-259.6	149.5	-203.7	-463.
Refrigeration and service industry machinery	205.2	220.6	15.4	195.8	-9.4	6.
Industrial machinery, nec...................	342.5	314.4	-28.1	312.1	-30.4	-58.
Electric distribution equipment.............	83.6	125.5	41.9	84.4	0.8	42.
Electrical industrial apparatus.............	158.6	170.6	12.0	180.5	21.9	33.
Household appliances........................	121.4	145.6	24.2	115.1	-6.3	17.
Electric lighting and wiring equipment......	180.1	235.2	55.1	170.4	-9.7	45.
Household audio and video equipment.........	84.7	13.1	-71.6	145.9	61.2	-10.
Communications equipment....................	266.9	148.7	-118.2	230.3	-36.6	-154.
Electronic components and accessories.......	585.0	416.2	-168.8	295.3	-289.7	-458.
Miscellaneous electrical equipment..........	155.7	114.9	-40.8	168.1	12.4	-28.
Motor vehicles and equipment................	972.9	1,037.9	65.0	1,049.0	76.1	141.
Aerospace...................................	549.7	551.8	2.1	675.4	125.7	127.
Ship and boat building and repairing........	164.6	286.0	121.4	177.9	13.3	134.
Railroad equipment..........................	37.6	67.3	29.7	49.5	11.9	41.
Miscellaneous transportation equipment......	74.4	65.2	-9.2	87.3	12.9	3.
Search and navigation equipment.............	165.7	128.4	-37.3	251.5	85.8	48.
Measuring and controlling devices...........	288.3	268.8	-19.5	251.0	-37.3	-56.
Medical equipment, instruments, and supplies	266.9	209.9	-57.0	183.8	-83.1	-140.
Ophthalmic goods............................	37.4	28.4	-9.0	38.5	1.1	-7.
Photographic equipment and supplies.........	86.1	91.1	5.0	107.8	21.7	26.
Watches, clocks, and parts..................	8.0	23.1	15.1	22.8	14.8	29.
Jewelry, silverware, and plated ware........	62.6	108.2	45.6	69.3	6.7	52.
Toys and sporting goods.....................	126.7	118.9	-7.8	134.2	7.5	-0.
Manufactured products, nec..................	240.0	275.2	35.2	256.3	16.3	51.
Meat products...............................	476.1	466.4	-9.7	495.4	19.3	9.
Dairy products..............................	149.2	181.0	31.8	141.9	-7.3	24.
Preserved fruits and vegetables.............	246.6	294.1	47.5	239.2	-7.4	40.
Grain mill products and fats and oils.......	160.4	159.2	-1.2	155.1	-5.3	-6.
Bakery products.............................	218.2	301.5	83.3	220.1	1.9	85.
Sugar and confectionery products............	101.9	115.3	13.4	110.9	9.0	22.
Beverages...................................	175.3	176.9	1.6	197.7	22.4	24.
Miscellaneous food and kindred products.....	184.4	180.4	-4.0	213.3	28.9	24.
Tobacco products............................	42.1	89.4	47.3	32.5	-9.6	37.
Weaving, finishing, yarn, and thread mills..	359.4	390.7	31.3	429.1	69.7	101.
Knitting mills..............................	195.0	210.5	15.5	212.5	17.5	33.
Carpets and rugs............................	63.3	47.8	-15.5	80.5	17.2	1.
Miscellaneous textile goods.................	53.3	50.3	-3.0	60.0	6.7	3.
Apparel.....................................	732.5	806.2	73.7	981.9	249.4	323.
Miscellaneous fabricated textile products...	238.3	232.1	-6.2	214.8	-23.5	-29.
Pulp, paper, and paperboard mills...........	227.6	224.5	-3.1	237.4	9.8	6.
Paperboard containers and boxes.............	221.0	230.9	9.9	258.3	37.3	47.
Converted paper products except containers..	246.3	240.4	-5.9	261.9	15.6	9.
Newspapers..................................	465.8	551.5	85.7	726.8	261.0	346.
Periodicals.................................	142.2	158.2	16.0	149.9	7.7	23.
Books.......................................	140.4	157.2	16.8	129.8	-10.6	6.
Miscellaneous publishing....................	92.2	92.9	0.7	50.2	-42.0	-41.
Commercial printing and business forms......	655.1	628.0	-27.1	543.5	-111.6	-138.
Greeting cards..............................	29.1	23.7	-5.4	29.2	0.1	-5.
Blankbooks and bookbinding..................	74.8	77.6	2.8	79.3	4.5	7.
Service industries for the printing trade...	59.4	56.1	-3.3	61.5	2.1	-1.

Industrial chemicals........................	266.5	255.8	-10.7	481.9	215.4	204.
Plastics materials and synthetics...........	159.4	158.0	-1.4	160.6	1.2	-0.
Drugs......................................	259.8	219.9	-39.9	238.8	-21.0	-60.
Soap, cleaners, and toilet goods............	155.1	150.8	-4.3	174.0	18.9	14.
Paints and allied products..................	55.3	61.7	6.4	58.3	3.0	9.
Agricultural chemicals......................	53.1	52.9	-0.2	60.7	7.6	7.
Miscellaneous chemical products............	92.9	101.7	8.8	89.1	-3.8	5.
Petroleum refining.........................	104.5	120.3	15.8	124.5	20.0	35.
Miscellaneous petroleum and coal products..	40.7	45.8	5.1	49.0	8.3	13.
Tires and inner tubes......................	80.2	102.6	22.4	87.3	7.1	29.
Rubber products, plastic hose and footwear.	189.5	198.2	8.7	205.3	15.8	24.
Miscellaneous plastics products, nec.......	714.1	680.5	-33.6	486.7	-227.4	-261.
Footwear, except rubber and plastic........	55.6	173.6	118.0	55.1	-0.5	117.
Luggage, handbags, and leather products,nec	54.0	109.7	55.7	69.3	15.3	71.
Railroad transportation....................	238.4	295.9	57.5	390.6	152.2	209.
Local and interurban passenger transit.....	469.1	706.8	237.7	331.8	-137.3	100.
Trucking and warehousing...................	1,861.2	1,790.2	-71.0	1,832.8	-28.4	-99.
Water transportation.......................	185.9	165.5	-20.4	263.2	77.3	56.
Air transportation.........................	1,074.4	807.7	-266.7	879.6	-194.8	-461.
Pipelines, except natural gas..............	15.1	17.0	1.9	16.3	1.2	3.
Passenger transportation arrangement.......	217.2	172.5	-44.7	133.6	-83.6	-128.
Miscellaneous transportation services......	203.3	195.0	-8.3	159.7	-43.6	-51.
Communications.............................	1,342.7	1,087.5	-255.2	1,372.5	29.8	-225.
Electric utilities.........................	496.3	465.8	-30.5	554.6	58.3	27.
Gas utilities..............................	188.1	223.4	35.3	410.8	222.7	258.
Water and sanitation.......................	240.6	234.2	-6.4	220.1	-20.5	-26.
Wholesale trade............................	6,733.8	5,882.9	-851.0	5,262.3	-1,471.5	-2,322.
Retail trade exc eating and drinking places	15,047.8	14,025.4	-1,022.4	14,672.4	-375.4	-1,397.
Eating and drinking places.................	7,587.2	7,535.7	-51.5	7,677.8	90.6	39.
Depository institutions....................	2,028.1	2,142.7	114.6	2,161.9	133.8	248.
Nondepository;holding & investment offices.	701.2	408.8	-292.4	829.9	128.7	-163.
Security and commodity brokers.............	614.4	350.1	-264.3	359.3	-255.1	-519.
Insurance carriers.........................	1,528.9	1,586.6	57.7	1,606.6	77.7	135.
Insurance agents, brokers, and service.....	854.5	879.1	24.6	1,218.5	364.0	388.
Real estate................................	1,744.8	1,728.6	-16.2	1,726.5	-18.3	-34.
Hotels and other lodging places............	1,726.1	1,773.8	47.7	2,245.8	519.7	567.
Laundry, cleaning, and shoe repair.........	544.0	732.3	188.3	565.6	21.6	209.
Personal services, nec.....................	344.4	341.1	-3.3	614.0	269.6	266.
Beauty and barber shops....................	802.1	1,004.0	201.9	759.0	-43.1	158.
Funeral service and crematories............	100.4	192.3	91.9	88.6	-11.8	80.
Advertising................................	265.9	261.6	-4.3	228.4	-37.5	-41.
Services to buildings......................	1,075.2	917.5	-157.7	816.6	-258.6	-416.
Miscellaneous equipment rental and leasing.	263.1	246.6	-16.5	228.6	-34.5	-51.
Personnel supply services..................	2,502.5	2,186.5	-316.0	1,258.4	-1,244.1	-1,560.
Computer and data processing services......	1,194.9	1,034.8	-160.1	357.6	-837.3	-997.
Miscellaneous business services............	2,232.9	2,059.8	-173.1	2,144.6	-88.3	-261.
Automotive rentals, without drivers........	184.1	157.7	-26.4	103.7	-80.4	-106.
Automobile parking, repair, and services...	1,145.0	1,019.4	-125.6	1,194.7	49.7	-75.
Electrical repair shops....................	146.3	151.7	5.4	174.9	28.6	34.
Watch, jewelry, & furniture repair.........	72.8	89.6	16.8	79.9	7.1	23.
Miscellaneous repair services..............	388.9	390.7	1.8	584.3	195.4	197.
Motion pictures............................	368.6	349.2	-19.4	272.4	-96.2	-115.
Video tape rental..........................	160.1	11.6	-148.5	0.0	-160.1	-308.

Producers, orchestras, and entertainers....	259.7	183.9	-75.8	206.6	-53.1	-128.
Bowling centers...........................	88.2	179.3	91.1	129.8	41.6	132.
Commercial sports.........................	*125.9*	*182.7*	*56.8*	*135.0*	9.1	*65.*
Amusement and recreation services, nec.....	1,139.3	830.8	-308.5	960.5	-178.8	-487.
Offices of health practitioners...........	2,958.2	2,579.9	-378.3	2,927.4	-30.8	-409.
Nursing and personal care facilities.......	1,696.4	1,258.4	-438.0	1,514.5	-181.9	-619.
Hospitals, private........................	3,780.1	3,323.7	-456.4	3,816.3	36.2	-420.
Health services, nec......................	1,205.7	531.2	-674.5	1,564.0	358.3	-316.
Legal services............................	1,158.4	1,183.4	25.0	1,330.6	172.2	197.
Educational services......................	2,079.2	2,552.0	472.8	1,552.9	-526.3	-53.
Individual & miscellaneous social services.	847.4	487.3	-360.1	843.6	-3.8	-363.
Job training and related services..........	304.1	353.7	49.6	390.1	86.0	135.
Child day care services....................	1,077.6	966.2	-111.4	1,072.9	-4.7	-116.
Residential care..........................	656.6	373.6	-283.0	657.0	0.4	-282.
Museums, botanical, zoological gardens.....	84.4	56.7	-27.7	73.3	-11.1	-38.
Membership organizations..................	2,145.9	1,769.6	-376.3	2,599.0	453.1	76.
Engineering and architectural services.....	870.8	951.3	80.5	547.7	-323.1	-242.
Research and testing services.............	583.8	448.7	-135.1	390.9	-192.9	-328.
Management and public relations...........	1,011.2	889.3	-121.9	395.0	-616.2	-738.
Accounting, auditing, and other services...	935.0	919.4	-15.6	1,221.5	286.5	270.
Private households........................	939.0	1,432.4	493.4	939.0	0.0	493.
U.S. Postal Service.......................	843.4	746.8	-96.6	796.1	-47.3	-143.
Federal electric utilities................	28.0	26.0	-2.0	40.2	12.2	10.
Federal government enterprises, nec........	127.6	109.3	-18.3	359.8	232.2	213.
Federal general government................	1,823.0	2,906.0	1,083.0	1,823.0	0.0	1,083.
Local government passenger transit.........	213.4	322.8	109.4	192.5	-20.9	88.
State and local electric utilities.........	86.2	80.1	-6.1	114.8	28.6	22.
State and local government enterprises, nec	589.4	498.3	-91.1	710.0	120.6	29.
State and local government hospitals.......	1,064.1	1,265.6	201.5	1,064.1	0.0	201.
State and local government education.......	8,524.6	10,281.7	1,757.1	8,524.6	0.0	1,757.
ate and local general government, nec....	6,006.3	6,542.0	535.7	6,006.3	0.0	535.

BUSINESS INVENTORY PRACTICES: Model and Analysis

Jay Berman, Bureau of Labor Statistics

I. Overview

The study of changes in business investment in inventories, which rarely exceed 1 percent of GDP, is often overlooked in favor of more marquee analyses. But, through improved inventory management, companies become more efficient and are therefore more responsive to changes in demand preferences and supply conditions. As a result, the importance of this area should not be ignored.

As part of the most recent U.S. economic and employment projections developed biennially by the Bureau of Labor Statistics[1], this paper introduces a model that simulates business inventory liquidation and accumulation practices by detailed industry.

Based on the model and prompted by claimed improvements in inventory management, this paper also quantifies the extent to which inventory investment has become more efficient and traces those benefits through the economy. This paper finds that without these efficiencies, during the 1991 U.S. recession, GDP might have declined by an additional $80 billion and might have caused a further reduction in employment of 1.4 million jobs.

II. The Model: Inventory levels by industry

The attention that researchers have given modeling inventory practices is mostly macro in nature. Typical studies, for example, are confined to addressing the interaction between total inventories and GDP or the determinate factors behind manufacturing inventory practices in aggregate. Detailed analysis of individual industry and commodity inventory trends have been largely ignored in favor of broader analysis.

As part of the Bureau of Labor Statistics' latest 1998 to 2008 projections of the US labor force, gross domestic product (GDP) and its components, industry output, and industry and occupational employment, data pertaining to historical inventory practices on behalf of individual industries have been developed. Quarterly inventory data from 1983 through 1997 are available for over 100 agricultural, manufacturing, transportation, and trade industries. These detailed industry inventory data were used to expand on previous, more general inventory studies through the formulation of a model of industry inventory accumulation and liquidation practices on a quarterly basis.[2]

The model's main structure is attributable to work done by Feldstein and Auerbach (Brookings, 1976), who hypothesized that a firm's inventory investment decisions are based on an educated sales expectation. Their target-adjustment model assumes that inventories adjust to a predetermined target level within one quarter while the target level itself responds more slowly. Firms anticipate change rather than assume the sales level will remain the same from one period to the next. The model assumes that only the portion of unanticipated sales that occurs late in the quarter will go uncorrected to any significant degree. In addition, F. Owen Irvine, Jr. (AER,1981) has developed a sales expectation formula, which found that retail inventory levels depend inversely on variations in estimated inventory carrying costs. Both the sales expectations and inventory carrying costs equations are incorporated into the BLS target-adjustment model.

The Feldstein--Auerbach target-adjustment model estimates industry inventory levels for finished goods. Inventories, however, are actually divided into three category types, representing different stages of fabrication: raw materials, work in progress, and finished goods. Industries, each with unique production processes, accumulate different types of inventories. For example, inventories held by the farm, wholesale and retail trade, and transportation industries are exclusively finished goods. The margin industries, who facilitate markets by bridging the gap between consumers and producers, do not produce goods; items sold by these industries are solely finished goods. In contrast, manufacturing industries predominately hold raw materials and work in progress inventories, finished goods inventories are relatively negligible. For instance, about 80 percent of the inventories historically held by the motor vehicles and equipment industry are raw materials and work in progress.

Since BLS analyzes the total economy, a model that encompasses each type of inventory is, therefore, required. Running different simulations by industry for each respective type of inventory and then summing the results was considered, but this was not

[1] For a detailed discussion of the Bureau's projections, see Norman C. Saunders and Betty W. Su, "The U.S. economy to 2008: a decade of continued growth," *Monthly Labor Review*, November 1999, pp. 5-18.

[2] For the model's supporting data, contact Jay Berman (202) 691-5692.

found to significantly improve the model's accuracy. Therefore, the model estimates total quarterly inventories by industry and does not differentiate between the three types of inventories. The specifications for this model follow.

Model specification

The target-adjustment model of inventory for finished goods first assumes that the stock of inventories will adjust within the quarter to the currently desired level, except for a small effect of unanticipated sales, implying that α is positive but quite small.

equation 1:

$$I_t = I_t^* + \alpha_0(S_t^e - S_t) + u_t$$

where

- I_t actual inventories of finished goods at the end of quarter t
- I_t^* desired inventories of finished goods at the end of quarter t
- S_t actual sales at end of quarter t
- S_t^e anticipated sales at end of quarter t

The firm's desired or target level of inventories is assumed to be influenced by a linear function of expected sales (S_t^e) and inventory carrying costs (C_t^e). It is assumed that each firm has a desired target level of inventory and that each firm, finding its actual inventory not equal to its optimum level, attempts only a partial adjustment towards the optimum level within any one period. The speed of adjustment is represented by the coefficient μ.

equation 2:

$$I_t^* - I_{t-1}^* = \mu(\alpha_1 + \alpha_2 S_t^e + \alpha_3 C_t^e - I_{t-1}^*) + \varepsilon_t$$

Unlike the stock-adjustment model, the sales forecast is not a naïve expectations assumption that the current level of sales will simply continue into the next quarter ($S_t^e = S_{t-1}^e$). A more proficient sales forecast formula, developed by Irvine (AER 1981), is last year's sales in the same month adjusted by the firm's recent sales experience. This formula adjusts a linear extrapolation by the amount such a linear extrapolation would have been off over the previous three months.

equation 2a:

$$S_t^e = S_{t-12}\left[\left(\frac{1}{3}\right)\left[\frac{S_{t-1}}{S_{t-13}} + \frac{S_{t-2}}{S_{t-14}} + \frac{S_{t-3}}{S_{t-15}}\right]\right]$$

The expected cost of holding inventory depends on the real interest rates and the relative price of the sector's good. The specification for the inventory capital cost measure is as follows:

equation 2b: $$C_t^e = \frac{P_t}{PC_t}(r_t - P_t^e)$$

where

- C inventory carry costs
- P_t retail price of the sector's goods
- r_t short-term interest rate
- P_t^e expected rate of inflation of the sector's goods over the inventory holding period.
- PC_t consumer price index

Solving for the level of desired inventories

equation 3:

$$I_t^* = (1-\mu)I_{t-1}^* + \mu\alpha_1 + \mu\alpha_2 S_t^e + \mu\alpha_3 C_t^e + \varepsilon_t$$

Substituting equation 3 into 1 yields

equation 4:

$$I_t = (1-\mu)I_{t-1}^* + \mu\alpha_1 + \mu\alpha_2 S_t^e + \mu\alpha_3 C_t^e + \alpha_0(S_t^e - S_t) + (u + \varepsilon_t)$$

To solve for I_{t-1}^*, which is not observable, use a lagged version of equation 1.

equation 5:

$$I_{t-1}^* = I_{t-1} - \alpha_0(S_{t-1}^e - S_{t-1}) - u_{t-1}$$

Combining equations 5 and 4, yields the final equation for the target-adjustment model of inventory accumulation:

equation 6:

$$I_t = (1-\mu)I_{t-1} - (1-\mu)\alpha_0(S_{t-1}^e - S_{t-1}) + \mu\alpha_1 + \mu\alpha_2 S_t^e + \mu\alpha_3 C_{t-1}^e + \nu_t$$

Per the above final functional form, the model includes four variables: lagged inventory levels, lagged sales anticipation error, sales expectations, and an inventory carrying cost measure. The sales anticipation error variable, which proved to be statistically insignificant, was dropped from the original specification. See Appendix A for each industry's regression coefficients and relevant statistical parameters.

Variable Clarification

Lagged quarterly inventory levels, by industry (I_{t-1}): Quarterly inventory levels by industry were derived by using two data sources: the Bureau of the Census' Annual Survey of Manufactures (ASM)[3] and the Bureau of Economic Analysis' National Income and

[3] For information pertaining to the Annual Survey of Manufacturers data, see "1996 Annual Survey of Manufactures." M96(AS)-1 (US Department of Commerce, Bureau of the Census, February 1998).

Product Accounts (NIPA)[4]. Because the majority of industries that accumulate and liquidate inventories are in the manufacturing sector, the industry breakdown offered by the ASM data proved critical. The NIPA data – quarterly inventory levels by major industry category – were used for all other industries that hold inventories.

The estimated speed of adjustment (μ), which is one minus the lagged inventory level coefficient, illustrates how quickly firms adjust their inventories to their targeted values. Following conventional theory, most of the manufacturing industries have relatively low estimated speeds of adjustment of approximately .25. This means that only 25 percent of the gap between the actual and the targeted inventory is eliminated within a quarter. On the other hand, the retail trade industry, which has different production dynamics, adjusts more rapidly with an average μ of .56.

Sales variables: (S_t) and (S^e_t): The Bureau of Economic Analysis' final sales data, which is GDP plus the change in inventory, was used in lieu of GDP. Because the NIPA final sales data are available only on a quarterly basis, the original formula specification was changed from monthly to quarterly.

In line with expectations, the retail trade industry's primary inventory level determinant is their expectation of future sales. This contrasts with the majority of manufacturing industries, both durable and non-durable, whose estimated coefficients indicate that last quarter's inventory level is the major inventory investment determinant. This is followed by an even split between their expectation of future sales and the cost restraints associated with holding inventories.

Lagged sales anticipation error ($S^e_{t-1} - S_{t-1}$): If the firm's sales expectation estimate differs from actual sales, inventories will be either accumulated or liquidated unexpectedly. To account for this, a lagged sales anticipation error variable was added. Note that the sooner firms are able to correct for this error, the smaller the estimated coefficient. Examining the regression results reveals that most of the industries readily correct for this error. In accordance with previous modeling work, this variable also commands a negligible role in determining inventory levels of most industries.

Inventory carrying cost measure (C_{t-1}): The inventory carrying cost measure, defined as the number of real dollars per year it costs to hold a unit of inventory, comprises two parts: the relative price of a sector's good and the real interest rate. Specifically, annual industry deflators were used as a price proxy for the retail price of sector's good (P_t). The GDP implicit price deflator (SA, 1992 =100) was utilized to define the price of all goods (PC_t). Based on the assumption that inventory is held for a relatively short period of time, the expected inflation rate of a sector's good also covers a short time horizon. Therefore, the expected rate of inflation for the sector's good (P^e_t) equals the actual rate of inflation observed over the previous two years. The prime bank interest rate, obtained from the Federal Reserve Bank, was used as a proxy for the nominal short term interest rate (r_t).

Therefore examined together, the formula states that inventory carrying costs increase when the relative price of the sector's good increases or when the real interest rate increases. A priori, negative inventory carrying cost coefficients were expected. However, after running the model, over half of the industries exhibited positive inventory cost coefficients. Explanations offered for this apparent anomaly include:

- The physical inventory facility needs of a particular industry are small. For example, the jewelry industry versus the furniture industry.
- Future price or sales expectations are very positive, therefore prompting increasing inventories regardless of relative cost.

Elaborating on this phenomenon, here is an example of two industries that have divergent inventory carry costs. First, the oil industry, which during the 1970's and early 1980's, experienced accelerated price increases due in part to the OPEC crisis. As a result, this industry had a positive incentive to hold inventories because tomorrow's market would bring forth higher prices. The industry's 1981 cost of holding a unit of inventory was a negative $58.93. On the other hand, the computer industry, given their rapid pace of technology improvements and product developments, face declining prices for their products. This industry's high rate of product obsoletion is therefore reflected in their relatively high cost of carrying inventories. For instance, the estimated cost of holding each unit of inventory for the computer industry was $96.83 in 1981. This illustrates the importance of the inverse relationship between an industry's expected cost of carrying inventories and their future price expectation, which is a function of the industry's observed price changes over the previous two years.

III. Estimating Inventory Change by Commodity

An integral part of the Bureau of Labor Statistics' projection process is the development of

[4] These accounts display the value and composition of national output and the distribution of incomes generated in its production. For more information, see "An Introduction to National Economic Accounting" (US Department of Commerce, National Technical Information Service, March 1985).

gross domestic product (GDP) or final demand estimates, which is GDP distributed among final users. The sources of demand that comprise GDP are categorized into four broad groups: personal consumption expenditures, business investment, foreign trade, and government purchases. GDP is a measure of the goods and services produced in the U.S. in a given year. When business production exceeds demand, inventories are accumulated and counted with that year's production. Likewise, when business production falls short of demand and past inventories are liquidated, those goods are subtracted from that year's production total because they represent production from a prior year. Because business investment includes changes in business inventories, the inventory levels by industry estimated by the model need to be converted to inventory changes by commodity.

This is accomplished within an input-output framework, which provides a snap-shot of all transactions within the economy at a given point in time and contains two main tables—a make table and a use table. The make table shows which commodities an industry produces or makes, while the use table shows the inputs required or used by an industry in producing those commodities. In order to yield the industry-to-commodity translation within the I-O system, the model's results—total inventory by industry—are allocated to the three types of inventories using a historical distribution. The level of finished goods by industry are then read through the make table, while raw materials and work in progress inventories are read through the use table. The results are added together to derive a total inventory level distribution by commodity.[5]

In order to derive a distribution of annual changes in business inventories, the present year's quarter four results are subtracted from the proceeding year's quarter four estimates. This method was employed to develop a reproducible and statistically viable annual time series of changes in business inventories by commodity.

IV. Inventory Change and Business Cycles: Behavior and Analysis

Improvements in inventory management have been expected as companies take advantage of technology and communication advances, just-in-time inventory systems, and more accurate sales forecast scenarios. Businesses continue to become more efficient and responsive to changes in demand preferences and supply conditions through enhanced inventory management. Prompted by these trends, this analysis quantifies the extent to which inventory investment has become more efficient and traces those benefits through the economy.

Inventory behavior and economic downturns: The often silent role that inventories play in our economy is examined using the historical data underlying the above model. In 1997, GDP amounted to $7.3 trillion, indicating daily production of about $28 billion. Inventory accumulation, hitting its historical pinnacle that year, amounted to only $63.2 billion—less than three day's production. However, the fact that inventory investment rarely exceeds 1 percent of GDP often masks its importance. In particular, the significant role of inventory management is brought to light when analyzing cyclical contractions in the economy.

Table 1, Part A traces the relationship between the peak-to-trough declines in GDP, final sales, and inventory investment during the last four recessions. The data show that changes in inventory investment consistently account for a major portion of recessionary declines in GDP. Specifically, during the last four recessions, inventory change has, on average, accounted for almost 50 percent of the peak-to-trough declines in GDP. Other researchers examining U.S. recessions prior to 1973 have found this to be even more apparent, averaging almost 100 percent.[6]

One conclusion to be drawn from the lead role inventories play during recessions is that declines in final sales are markedly less volatile than declines in GDP. The wane in the amount of goods and services demanded by the economy has historically been more benign relative to the amount supplied by businesses. The mismatch of demand and production is then absorbed by inventories.

A catalyst behind this phenomenon might be a misperception by industries of demand volatility. During an economic downturn, businesses choose to err on the side of caution by cutting production and relying on inventories to meet potential shortfalls.

Accentuating this point is Table 1, Part B, which presents quarter-to-quarter movements during each recession, plus three quarters following each trough. In particular, during some recessionary quarters, the decline in inventory investment has been greater than the decline in GDP—indicating that final sales have actually increased during these periods of recession.

Analysis: As discussed earlier, the difference between the change in GDP and changes in final sales is

[5] By using the use table to translate inventory levels by industry to commodities, the model assumes that every commodity each industry uses is being accounted for and they are represented in their correct proportions.

[6] This is a continuation of the study by Alan S. Binder, "Inventories and the Structure of Macro Models," *AEA Papers and Proceedings*, May 1981.

inventory change. Congruently, the difference between what companies produce and what they sell is absorbed by the liquidation or accumulation of inventories. Therefore, as industries refine their inventory investment behavior, the gap or ratio between the change in GDP and final sales should narrow. Minimizing the reliance on inventory and, consequently, narrowing the gap between what the economy produces and what it sells should result from improved inventory investment behavior.

The resulting hypothesis is that over time, inventory change should steadily contribute less toward a recession's severity. The economy might experience a downturn, but the acuteness of the decline is enhanced or mitigated by inventory investment practices. Therefore, as inventory management practices improve, recessions should becoming relatively less severe—as reflected in a narrowing of the gap between the change in GDP and final sales. The ideal situation would be a one-to-one ratio of the change in GDP to final sales. In such cases, businesses perfectly gauge the decline in final sales and reduce their production in line with the change in demand, thus mitigating the trough.

It is important to clarify the premise that recessions should become relatively less severe in step with improved inventory investment practices. This analysis was not concerned with whether recent recessions have been less severe relative to previous downturns; in fact, Table 1 shows, the decline in GDP during the 1990-91 recession was greater than that experienced in 1980. Rather, the reach of this study was to examine inventory's contribution to a given recession and to test the extent by which improved inventory practices mitigated an individual recession. This was then juxtaposed against a theoretical scenario, which assumed that these improvements did not exist. This study, focusing on recessionary periods in which the importance of inventory management is underscored, provides quantitative estimates of the impact of improved inventory holding practices on specific recessions.

Table 2 illustrates that during the last four recessions, the ratio between the change in GDP and final sales has steadily declined from 3.26 percent during the 1974 recession to 2 percent during the most recent recession of 1990-91. As the decline in this ratio illustrates, the impact of better inventory management on the economy is striking. For example, had the U.S. economy in 1990 experienced the same GDP-to-final sales ratio it did in 1973, GDP would have contracted by about $203 billion, or 3.3 percent, instead of $124 billion, or 2.0 percent.

The additional $80 billion decline in GDP would have caused a further reduction of 1.4 million jobs for a total decline of over 4 million jobs.[7] (See Table 3.) The severity of the 1990 recession would have almost doubled if advancements in inventory management had stagnated at the level existing in the early 1970's.

The number and types of jobs affected by this scenario were estimated using an input-output system that traces a given level of demand through the production chain. Using this structure, the employment in each industry, including the industries that supply inputs to the production process, can be determined. Table 3 highlights the top 10 industries most affected by this scenario. If industries err in their decision to cut production, employment in the wholesale trade industry, which sells merchandise to retailers and industrial users, experiences the greatest decline of about 295 thousand workers. The relatively large hypothetical drop in agricultural industry employment—135 thousand additional workers—points to the strides this sector has made to enhance its inventory practices and meet changing market conditions. The household furniture industry is another example of an industry taking advantage of technology and improved management practices. Industry employment would have declined by an additional 35 thousand jobs or 12 percent.

V. Concluding Remarks

This analysis illustrates that qualitative analysis can emerge from examining inventory measures with a more unorthodox, micro perspective. As part of the biennial projections process of the Bureau of Labor Statistics, a working methodology for compiling annual inventory data by industry and commodity has been achieved. Specifically, both Annual Survey of Manufacturing and National Income and Product Account data was used to compile inventory data by industry. The translation of inventory estimates by industry to the detailed commodities being liquidated and accumulated was then realized using an input-output accounting system. It is hoped that a void has been filed in this arena and researchers will use this system to extend their analysis beyond existing macro studies.

Following this premise, a statistically viable econometric model was assembled for projecting inventory levels by detailed industries. In addition, this system was also used to ascertain the effects that improved inventory management has had on the U.S. economy. It was determined that without these efficiencies, during the 1991 U.S. recession, GDP might have declined by an additional $80 billion and

[7] For more information on how the transition from production to employment was made, see "BLS Handbook of Methods" (US Department of Labor, Bureau of Labor Statistics, April 1997).

might have caused a further reduction in employment, especially in the wholesale trade, agriculture, trucking and courier, household furniture, and electronic components industries.

Given the continued aggressive pace of technological advances and innovative ways companies conduct business, the important role that business inventories play in the economy should continue, providing an important area for future inquiry.

References

Binder, A. S. (1981). Retail Inventory Behavior and Business Fluctuations, *Brookings Papers on Economic Activity*, 2, 443-520.

Binder, A. S. (1981). Inventories and the Structure of Macro Models. *AEA Papers and Proceedings*, 71 (2), 11-16.

Feldstein, M & Auerbach, A. (1976). Inventory Behavior in Durable-Goods Manufacturing: The target-Adjustment Model, *Brookings Papers on Economic Activity*, 2, 351-393.

Irvine, Jr., F.O. (1981). Retail Inventory investment and the Cost of Capital, *The American Economic Review*, 71 (4), 633-647.

Lovell, M. (1961). Manufacturers' Inventories, Sales Expectations, and the Acceleration Principle, *Econometrica*, 29 (3), 293-314.

MODELING THE DEMAND FOR SKILLS

Charles Bowman
U.S. Bureau of Labor Statistics

The widespread industrial restructuring of the past two decades combined with the more recent phenomenon of very tight labor markets has focused new attention on the availability and quality of labor resources in the United States. A well-trained and flexible workforce is widely seen as a key to higher overall living standards and a more equitable distribution of income. It is not clear, however, which policies and programs would be most effective in bringing these outcomes about. Not all training contributes to marketable skills and not all skills are in short supply. Clearly, there is a need for much more information on what skills will be needed in the future and on how training can best be structured to develop these skills.

Aside from case studies, most labor market projections that differentiate among types of labor have relied on some form of occupational analysis. The Bureau of Labor Statistics (BLS), for example, uses an input-output based model of occupational demand to address a wide variety of public policy issues and to develop occupational forecasts and related counseling materials for those planning careers and seeking jobs. Although this approach continues to be useful and has provided many insights into the evolution of U.S. labor markets, some current issues cannot be fully addressed in terms of occupational change alone. Perhaps the leading issue of this type is that of the changing requirements for and interrelationships among education, training, skills and jobs. The challenge for the future is to find ways to integrate these additional dimensions of labor input into forecasting models.

We begin with some general comments on the meaning and measurement of job skills. We then turn to the BLS employment model and its underlying database to outline the broad patterns of occupational change in the United States, both historically and as forecasted over the decade ahead. We look first at changes in the occupational structure itself and then at the implications of those changes for the training and education levels of the workforce. Our objective is to see whether we can discern any evidence of skill change in these patterns. Next, we present an industry-level index of skill change computed from the historical data underlying the BLS employment model. While based on somewhat restrictive assumptions, the index provides a more precise and quantitative assessment of skill change than afforded by analysis of changes in education and training requirements. Finally, we discuss recent developments in U.S. statistical programs that promise to overcome some of the limitations of past efforts and to open up new possibilities for labor modeling.

Measures of skill

There are at least two basic ways of defining skills in the labor force. One focuses primarily on the individual while the other takes the job itself as its focus. The labor composition index, produced by the BLS as part of its productivity measurement program, is a relatively sophisticated example of the former approach (U.S. Bureau of Labor Statistics 1993). Essentially, a quality-adjusted labor input measure is constructed by weighting the hours worked by each sex/experience/education cell by its wage. The difference between the adjusted labor index and a conventional index based simply on hours can then be interpreted as a measure of skill change. In this view, individuals accumulate a store of intellectual capital over time through education, formal training and learning by doing and this in turn raises their value in the labor market. Measures of this sort generally show a significant rise in skill levels over the past 2 or 3 decades since the underlying determinants, educational attainment and accumulated years of work experience, both exhibit a strong upward secular trend.

From the point of view of labor market analysis this approach has several weaknesses. First, there is no necessary connection between the accumulated intellectual capital of the workforce and the actual requirements of the job market. Research by BLS, for example, has consistently shown that about one quarter of college graduates occupy jobs for which a bachelor's degree is neither necessary nor usual (Mittlehauser 1998). Second, human capital measures generally have limited or no specificity with respect to particular skills. Third, there is no way to relate these measures to detailed industries and/or occupations that are often the focus of labor market policy initiatives.

The second way of looking at skills and the one adopted here is to focus on the requirements of specific jobs or, more precisely, occupations. Occupational analysis focuses attention on highly specific jobs and skill requirements. Within this general orientation, Spenner (1985) suggests that there are three strategies for assessing skill change: non-measurement, indirect measurement and direct measurement. Non-measurement involves making inferences, generally qualitative, based on such things as ratios between blue-collar workers and professionals or production workers versus non-production workers. The discussion below concerning the changing occupational structure of the United States and its education and training consequences is an example of this approach. Indirect measurement utilizes things like wages or education as proxies for skills. An industry-level estimate of skill change based on this strategy is also presented below. Direct measurement involves analysis in terms of the specific skills or skill sets associated with jobs such as substantive complexity or autonomy. While this strategy has been employed frequently in the past and has generally been seen as the most promising one its applicability is seriously limited by the inadequacy of current data on the skill content of jobs. In the final section we discuss a new data collection initiative that promises to expand greatly the possibilities of this approach.

Occupational trends

Over the decade from 1988 to 1998 the U.S. economy added over 20 million net new jobs (Table 1). Corresponding growth among the major occupational groups and within industries was by no means uniform. Professional specialty occupations grew the fastest of all groups and also added the largest number of new jobs, approximately 4.8 million. This group includes a wide variety of generally high-paying and skill-intensive jobs ranging from physicians, engineers and scientists to artists and entertainers. Not surprisingly, growth was concentrated in the rapidly growing services industries such as health care, education and business services. Nonetheless, these occupations also expanded rapidly in areas like manufacturing which showed little or no overall employment growth over the period. This reflects in large part the widespread adoption of computer technology and technologically advanced manufacturing methods that in turn require large numbers of engineers, systems analysts and similarly high-skilled occupations.

At the other end of the spectrum the service occupations group added the second largest number of new jobs, nearly 4 million, over the 1988-1998 period. These jobs, more often than not, are low paying with modest skill and education requirements. Government, health care and the retail trade sector, which in the U.S. industrial classification includes restaurants, accounted for about half of such workers. Other demands came from protective service industries, cleaning and janitorial services and a wide variety of personal service providers. As with professional workers the growth of these occupations was mediated by the industrial restructuring of the U.S. economy that was in full swing in this period. Unlike professional occupations, however, this group as a whole does not appear to have been influenced to a substantial degree by technological change although the widespread use of computers has certainly changed the nature of these jobs to some degree.

Several of the major occupational groups gained jobs over the decade but at rates far below the growth of employment as a whole. These include administrative support and clerical occupations, precision production and craft workers, machine and plant operators and agricultural workers. Many of these types of jobs are primarily located in manufacturing industries. With little or no employment growth in most manufacturing industries demand for many of these occupations is likely to be limited to replacement of existing workers. As in other occupational areas, the expanding service economy was the driving force behind what growth did occur.

Table 1 also contains a forecast of occupational trends generated by the BLS employment projections system for the 1998-2008 period. As illustrated in Chart 1, the projections system consists of a conventional input-output based model of industrial activity augmented by relatively detailed labor supply and occupational demand components. The occupational demand component itself consists of an industry-occupation matrix, the columns of which describe the occupational input structure of each of the 262 industries in the system. A consistent annual series of these matrices has been developed for the 1983-1998 period (U.S. Bureau of Labor Statistics 2000). The matrices are based on establishment surveys and are designed to be as compatible as possible with the input-output based industry component.

In developing occupational projections, analysts make explicit forecasts of the industry-occupation coefficients based on a wide variety of occupation-specific information including any trends observed in the coefficients themselves. The final forecast of occupational employment, of course, depends not only on these coefficients but also on changes in industrial structure and productivity that arise elsewhere in the system.

While there are exceptions, the forecast for the most part continues the trends observed over the preceding decade. There are no signs in the data so far to indicate that the industrial and occupational changes observed over the past decade or two are abating. Thus, employment growth in professional specialty occupations is expected to continue to lead both relatively and in terms of the number of new jobs.

Table 2 offers a considerably more detailed view of the likely patterns of occupational change over the next decade. The thirty occupations listed in the table are those with the largest expected job growth and account for about half of the total net change in employment forecasted over the 1998-2008 period. Most of these occupations are concentrated in the four industry sectors expected to dominate job growth: retail trade, business services, health care and education. Not surprisingly, all of the major computer-related occupations are on the list accounting for about 1.5 million new jobs. There are nearly a million new jobs in nursing occupations and a similar number in retailing including food service.

Table 2 also contains information on the relative earnings and typical education and training requirements associated with these occupations. What is most striking about this aspect of the table is the broad range of both education and training requirements and income potential exhibited by these large-growth occupations. In particular, it is difficult to see in these data any clear-cut skill bias in either direction. In the following section, however, we look in more detail at the education and training implications of occupational change and try to develop a more focused view of changing skill requirements from this perspective.

Education and training

In the BLS model each of the more than 550 detailed occupations is linked to one of eleven education and training categories (Wash 1995-96). While a variety of data sources enter into making each assignment, they are ultimately dependent on the judgment of occupational specialists. For each occupation, that category of education and training is selected which best reflects the manner in which most workers become proficient in their job. This includes both the mental and physical requirements of the job as well as employer preferences. Where an occupation exhibits multiple entry paths a decision is made as to which of them is in some sense the preferred or typical one. Over time, of course, education and training requirements of specific occupations can and do change. For this reason, the assignments are reviewed by occupational specialists every other year and updated when necessary.

Table 3 contains a tabulation of wage and salary occupational employment in terms of the 1998 education and training category assignments. Skill upgrading in terms of increased education and training requirements within detailed occupations is therefore ruled out by definition. Since the classification system has only been in place since 1994 there is little empirical evidence yet as to how important such upgrading may be. However, given the relatively narrow occupational categories this is unlikely to be a major factor over short to medium-length periods.

Given the assumption of fixed requirements, the changes shown in Table 3 can be interpreted the education and training consequences of shifts in the occupational structure. At most there is evidence of a slight overall shift toward occupations with higher education requirements but even this is concentrated in the earlier part of the period. Since 1992 occupations requiring a bachelor's degree or higher have accounted for 20.8 percent of employment while the proportion requiring only a bachelor's degree has remained at around 11.8 percent. On the other hand, jobs requiring associate degrees, generally, two years of post-secondary education, do show a steady rise. These jobs tend to be technical in nature and are concentrated in the health care and computer fields.

At the other end of the spectrum there is some evidence of reduced requirements among jobs requiring no specific education or training beyond the secondary school level. Here we see a steady decline in jobs requiring moderate to extensive on-the-job training and a concomitant increase in jobs requiring on-the-job training of a month or less.

Table 4 shows the education and training implications of BLS' most recent occupational forecasts. The results are based on the same set of 1998 education and training category assignments used to develop the data in table 3. The table shows that nearly 55 percent of expected job openings have no post-secondary educational requirements and require one year or less of on-the-job training. Most of these jobs, in fact, have training requirements of only a month or less. Table 4 also shows the income distribution within each of the education and training categories. These data imply that on average there is a large positive return to education and training but also that a large portion of future job openings will be in jobs that historically have paid relatively low wages.

To the degree that education and training requirements can be taken as a rough indicator of changes in skill requirements, there is very little evidence of skills upgrading over all. At most, there appears to be a slight shift toward jobs requiring at least some college training offset by a decided shift toward jobs with the most limited requirements.

Industry skill requirements

This section presents an indirect index of skill change computed on an industry-by-industry basis.[1] The index is indirect because the economy-wide relative wage of each occupation is used as a proxy for its skill level. Further, the relative wage proxy is fixed at its 1998 level. Thus, as with the education and training measure, we rule out skill shifts within narrowly defined occupations. This appears to be reasonable given the high degree of occupational specificity used in constructing the index and the rigidity of occupational classification systems. More troubling is the assumption that the 1998 relative wage of an occupation can serve as a proxy for its skill level. One question we would like to answer in this regard is whether the occupational wage distribution is stable. Unfortunately, occupational wage data at the level of detail needed to construct the type of index reported here has only been available since 1997. Tests do show, however, that using 1997 weights would make no appreciable difference in the results.

Another potential limitation is that even if the wage distribution is stable it may not be systematically related to skill differentials. Howell and Wolff (1991), for example, report poor correlation between earnings and a direct measure of skill. Lacking a reliable direct measure of skill compatible with the occupational data used here, there is no way to resolve this question definitively. Consequently, the measure of skill presented here has to be taken as tentative with due regard to the assumptions on which it is based. As suggested below, however, improvements in occupational data may eventually allow us to address this question.

The skill change index represents the percent change in an industry's wage bill due solely to changes in the occupational structure of that industry. An increase, for example, indicates that the industry has moved to a higher wage input structure, given the relative occupational wage structure of 1998. Taking the assumptions noted above, this may be interpreted as an increase in industry skill requirements.

The calculations are based on the time-series of industry-occupation matrices discussed above which were developed as part of the BLS employment projections system. The occupational wage data for 1998 are based on the BLS' Occupational Employment Survey. The survey was expanded in 1997 to include occupational wage data for the first time. As a result the occupational employment and wage data underlying the indirect skill index could be derived from the same establishment survey. Prior to 1997 occupational wage data had to be based on household surveys, introducing major problems of comparability.

While there are exceptions most industries exhibit a relatively small positive or negative change in skill requirements. There are significant increases in a number of manufacturing industries: computer manufacturing, publishing, apparel and guided missiles and space vehicles. A number of finance and insurance and transportation industries also show gains. Overall, however, there is little evidence of a pervasive change in skill requirements. (Complete industry results are available from the author on request).

Table 5 presents a summary of skill change for major sectors. The summary measures are calculated as 1998 employment-weighted averages of the industry data. As such they are interpreted in essentially the same way as the detailed measures. Based on the data in this table there appears to be little difference between goods-producing and services-producing industries. Most of the sectors show positive but small increases in skill levels overall. The exceptions are the Mining and construction sector that shows a decline over the latter part of the period and the Trade sector that declines over the whole period.

The estimates of skill change shown in table 5 include only intraindustry effects. The skill change measure was also calculated for the economy as a whole. While the interpretation of the result is the same, skill change now refers to reallocation of labor inputs throughout the economy. The difference between this and the average intraindustry effect provides an estimate of the amount of skill change attributable to interindustry employment shifts. These results are shown in Table 6. In all but one of the sub-periods the interindustry skill effect enhances the intraindustry effect but the overall effect is still relatively small.

In general, the results discussed in this section suggest that there has been a positive but small increase in skill requirements over the 1988-1998 period. Because the results depend on an indirect measure of skill change they are dependent on a number of restrictive assumptions and cannot be taken as definitive. On the other hand, the finding of a slow increase in skill requirements over time is generally consistent with a

[1] The measure developed here is similar to one proposed by Murphy and Welch (1993).

number of studies for earlier periods utilizing both direct and indirect measures (for example, Spenner(1983), Murphy and Welch (1993), Rumberger(1981).

Given the inherent limitations of indirect methods, improvements in our understanding of skill change in the work force will no doubt require improved direct measurement approaches. In the next section we discuss some data issues related to the feasibility of developing such measures.

Advances in labor market information

Most direct measures of skill change have utilized the U.S. Department of Labor's Dictionary of Occupational Titles (DOT) as the basic source of information on the skill content of jobs. The DOT, which was developed in the late 1930's to aid employment counselors and others involved in job placement, contains detailed information on nearly 13,000 jobs. The DOT describes the tasks of each occupation in terms of a functional relationship to data, people and things with scales to indicate the complexity of the relationship. The latest edition also contains occupation-specific information on a large number of variables such as training time, working conditions and physical and mental demands.

In spite of the wealth of information contained in the DOT, it has become less useful over time for analyzing skill change in the U.S. economy. Partly this is due to the inherent difficulty of keeping a database of this sort current. Most of the occupations in the DOT were last updated in 1977. Added to this is a bias toward manufacturing occupations, a reflection of the industrial structure and employment situation of the 1930's when the DOT was designed. Besides the question of currency, the structure of the DOT makes it less than ideal as a vehicle for studying labor market skills. First, it is a task-based system. It focuses on how tasks are carried out rather than what abilities are needed to accomplish those tasks. Second, it is based on an obsolete system of occupational classification that does not reflect the modern economy and is therefore difficult to link to related sources of labor market information. Third, it offers no easy way to compare requirements across occupations. Fourth, it is not sample-based and there is no way to gauge how representative it is of actual occupational requirements. Finally, coverage of education and training related to occupations is very limited.

Given the severity of the problems facing the DOT a decision was reached in the 1990's to completely redesign it. What emerged was the *Occupational Information Network,* or O*NET for short. The O*NET system is designed to greatly improve upon the content and usefulness of the DOT. It is intended to serve the multiple needs of job seekers, researchers and policy makers.

The organizational framework of the new system is the content model consisting of six domains in which information on each occupation is grouped. The Worker Characteristics domain contains information on abilities, values and interests, and work styles. These are seen as reflecting relatively enduring characteristics of individuals that can influence job performance. The Worker Requirements domain deals with an occupation's need for general skills, knowledge and education. Skills are further subdivided into basic skills such as reading, communication and critical thinking and cross-functional skills such as problem-solving, social and technological skills. The Experience Requirements domain contains information on the experience needed to perform in a job. Experience requirements are defined in terms of categories called job zones that are similar to the education and training categories used in the BLS model. This domain will also include links to licensure requirements.

The data contained in the Worker Characteristics, Worker Requirements and Experience Requirements domains are worker-oriented and together describe the demands placed on individuals. The remaining domains are work-oriented and describe the nature of the work itself. Labor Marker Requirements provides links to related data about occupations such as wages and BLS employment projections. The Occupational Requirements domain deals primarily with the work activities that make up a job, the environment in which the job is done and its organizational context. The final domain, Occupation-Specific Requirements differs from the others in that the variables that comprise it may be different for each occupation. In the case of the other five domains the same set of variables is used to describe each occupation.

At present the O*NET content model is populated primarily by data adapted from the DOT. The nearly 13,000 DOT occupations have been replaced by about 1000 categories based on the latest Standard Occupational Classification. Data collection is set to begin in late 2000. The goal is to collect data on the hundreds of descriptors and associated scales that describe each occupation. Current plans are to survey about one-third of the roughly 1000 occupations in each of the next three years. In general, the respondents will be incumbents in the occupation, selected by means of a probability sample of establishments. Each respondent will complete one of four questionnaires dealing with skills, work context, knowledge or generalized work activities. Demographic information about the respondent will also be collected along with a description of the tasks involved in the occupation. The results will be used to populate the various dimensions of the content model for each surveyed occupation.

The O*NET system addresses most of the shortcomings of the DOT and, if data collection proceeds as planned, it will offer an unprecedented insight into the skill composition of the U.S. workforce. It will take time, however, to fully realize it's potential. Over the course of the initial 3-year data collection period the results will be continuously analyzed and it is likely that survey methods and other collection parameters will change. Nonetheless, within a year the program should begin producing data on skills and other occupational characteristics which goes far beyond anything now available.

Conclusions

The empirical evidence presented in this paper generally supports the view that there has been at most a small increase in skill requirements over the past decade and that this is likely to hold true over the next ten years. Analysis of occupational and educational trends shows that while professional and technical jobs with relatively extensive educational requirements are growing the fastest large numbers of jobs at very low skill levels are also being created. The indirect measure of skill change presented in this paper supports this conclusion in that most industries show little evidence of upgrading and the overall change in skill requirements while positive is quite small. While convincing, none of the evidence presented here measures skill directly. Such measures are difficult to construct because of severe data limitations. The O*NET data collection program promises to remedy this and should allow us to develop much more precise estimates of skill change in the future.

References

Howell, David R., and Edward N. Wolff (1991). "Trends in the Growth and Distribution of Skills in the U.S. Workplace, 1960-1985." *Industrial and Labor Relations Review* 44(3): 486-502.

Mittlehauser, Mark (1998). "The Outlook for College Graduates, 1996-2006: Prepare Yourself." *Occupational Outlook Quarterly* Summer: 2-9.

Murphy, Kevin M. and Finis Welch (1993). "Occupational Change and the Demand for Skill, 1940-1990." *American Economic Association Papers and Proceedings* 83(2):122-126.

National Academy of Sciences (1999). *The Changing Nature of Work: Implications for Occupational Analysis* (Washington: National Academy Press).

Rumberger, Russell (1981). "The Changing skill Requirements of Jobs in the U.S. Economy." Industrial and Labor Relations Review 34(4): 578-590.

Spenner, Kenneth I. (1985). "The Upgrading and Downgrading of Occupations: Issues, Evidence, and Implications for Education." *Review of Educational Research* 55(2): 125-154.

Spenner, Kenneth I. (1983). "Deciphering Prometheus: Temporal Change in the Skill Level of Work." American Sociological Review 48(December): 824-837.

U.S. Bureau of Labor Statistics (2000). *The National Industry-Occupation Employment Matrix: 1983-98 Time Series Documentation.*

U.S. Bureau of Labor Statistics (1993). *Labor Composition and U.S. Productivity Growth, 1948-90 (BLS Bulletin 2426).*

Wash, Darrel Patrick (1995-96). "A New Way to Classify Occupations by Education and Training." *Occupational Outlook Quarterly* Winter: 29-40.

Chart 1. The BLS Employment Projections System

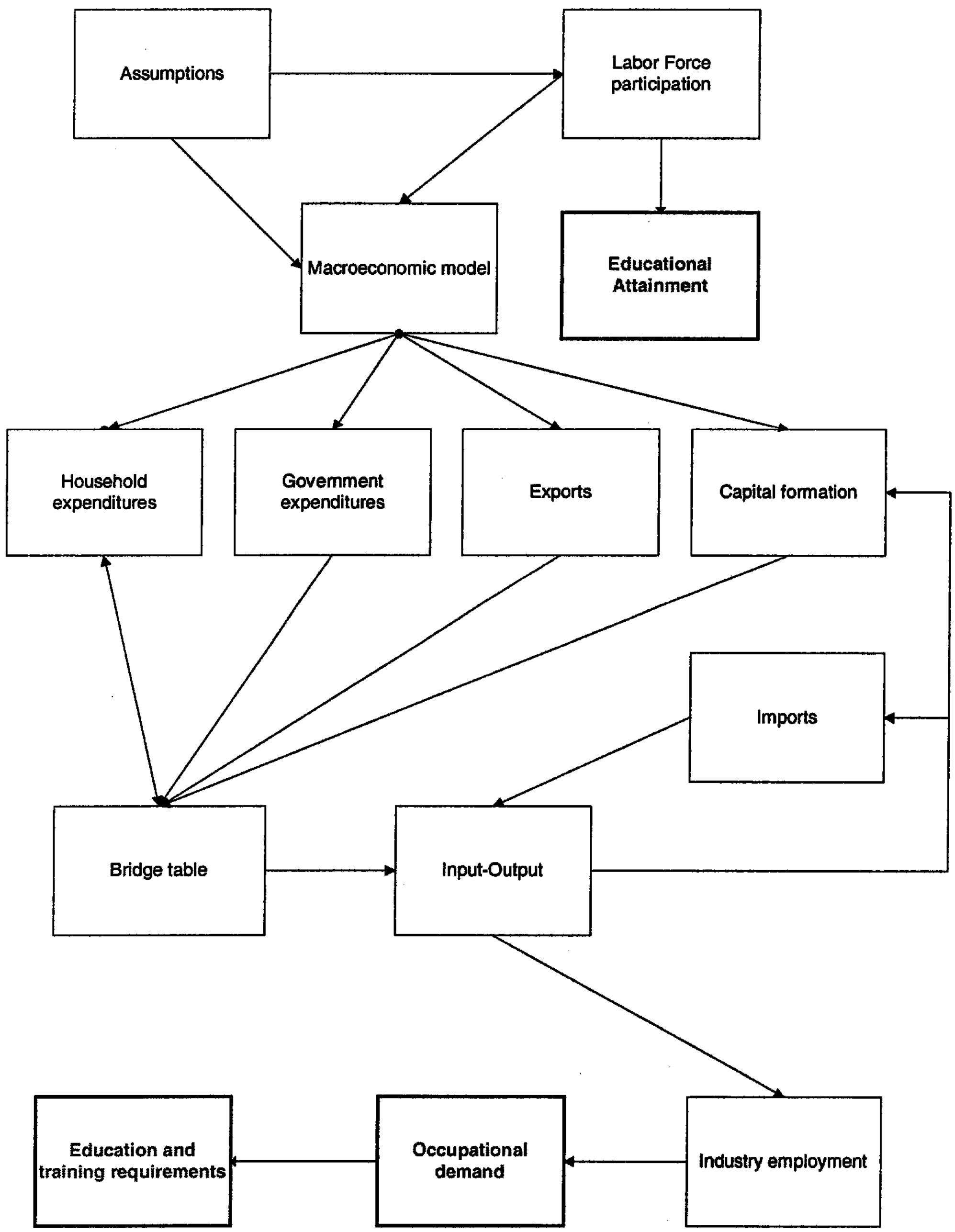


Table 1. Employment by major occupational group, 1993-1998 and projected to 2008

Occupational group	Employment			Employment Change		Percent Distribution		
	1988	1998	2008	1988-1998	1998-2008	1988	1998	2008
Total	120010	140514	160795	20504	20281	100.0	100.0	100.0
Executive, administrative and managerial	12330	14770	17196	2440	2426	10.3	10.5	10.7
Professional specialty	15035	19802	25145	4767	5343	12.5	14.1	15.6
Technicians and related support	3880	4949	6048	1069	1099	3.2	3.5	3.8
Marketing and sales	12390	15341	17627	2951	2286	10.3	10.9	11.0
Administrative support including clerical	22251	24461	26659	2210	2198	18.5	17.4	16.6
Service	18554	22548	26401	3994	3853	15.5	16.0	16.4
Agriculture, forestry and fishing	4224	4435	4506	211	71	3.5	3.2	2.8
Precision production, craft and repair	14333	15619	16871	1286	1252	11.9	11.1	10.5
Operators, fabricators and laborers	17012	18588	20341	1576	1753	14.2	13.2	12.7

Table 2. Occupations with the largest job growth, 1998-2008
(Employment in thousands of jobs)

Occupation	Employment 1998	Employment 2008	Change 1998-2008 Number	Change 1998-2008 Percent	Quartile rank by 1997 median hourly earnings*	Education and training category
Systems analysts	617	1194	577	94	1	Bachelor's degree
Retail salespersons	4056	4620	563	14	4	Short-term on-the-job training
Cashiers	3198	3,754	556	17	4	Short-term on-the-job training
General managers and top executives	3362	3913	551	16	1	Experience plus bachelor's
Truck drivers light and heavy	2970	3463	493	17	2	Short-term on-the-job training
Office clerks, general	3021	3484	463	15	3	Short-term on-the-job training
Registered nurses	2079	2,530	451	22	1	Associate degree
Computer support specialists	429	869	439	102	1	Associate degree
Personal care and home health aides	746	1179	433	58	4	Short-term on-the-job training
Teacher assistants	1192	1567	375	31	4	Short-term on-the-job training
Janitors and cleaners, including maids and housekeeping cleaners	3184	3549	365	11	4	Short-term on-the-job training
Nursing aides, orderlies, and attendants	1367	1692	325	24	4	Short-term on-the-job training
Computer engineers	299	622	323	108	1	Bachelor's degree
Teachers, secondary school	1426	1749	322	23	1	Bachelor's degree
Office and administrative support supervisors and managers	1611	1924	313	19	2	Experience in a related occupation
Receptionists and information clerks	1293	1599	305	24	3	Short-term on-the-job training
Waiters and waitresses	2019	2322	303	15	4	Short-term on-the-job training
Guards	1027	1321	294	29	4	Short-term on-the-job training
Marketing and sales worker supervisors	2584	2847	263	10	2	Experience in a related occupation
Food counter, fountain, and related workers	2025	2272	247	12	4	Short-term on-the-job training
Child care workers	905	1141	236	26	4	Short-term on-the-job training
Laborers, landscaping and groundskeeping	1130	1364	234	21	3	Short-term on-the-job training
Social workers	604	822	218	36	2	Bachelor's degree
Hand packers and packagers	984	1197	213	22	4	Short-term on-the-job training
Teachers, elementary school	1754	1959	205	12	1	Bachelor's degree
Blue-collar worker supervisors	2198	2394	196	9	1	Experience in a related occupation
College and university faculty	865	1061	195	23	1	Doctoral degree
Computer programmers	648	839	191	30	1	Bachelor's degree
Adjustment clerks	479	642	163	34	3	Short-term on-the-job training
Correctional officers	383	532	148	39	2	Long-term on-the-job training

* 1=very high ($16.25 and over), 2=high ($10.89 to $16.14), 3=low ($7.78 to $10.88), and 4=very low (up to $7.76).

Table 3. Wage and salary employment by education and training category, 1986-1998

Education and training category	Percent distribution							Change is share
	1986	1988	1990	1992	1994	1996	1998	1986-1998
Total, all occupations	100.0	100.0	100.0	100.0	100.0	100.0	100.0	---
Bachelor's degree and above	19.5	20.2	20.4	20.9	20.8	20.9	20.8	1.3
First professional degree	1.2	1.1	1.2	1.2	1.2	1.2	1.1	-0.1
Doctoral degree	0.9	0.9	0.8	0.9	0.8	0.8	0.8	-0.1
Master's degree	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.1
Work experience, plus a degree	5.8	6.3	6.2	6.1	6.1	6.2	6.2	0.4
Bachelor's degree	11.0	11.2	11.5	11.8	11.8	11.8	11.9	0.9
Postsecondary education and training	6.5	6.5	6.6	6.8	6.7	6.7	6.6	0.2
Associate degree	3.1	3.2	3.3	3.5	3.5	3.6	3.9	0.7
Postsecondary vocational training	3.4	3.3	3.2	3.3	3.1	3.1	2.8	-0.6
On-the-job training (OJT) or experience	74.0	73.3	73.0	72.4	72.6	72.4	72.5	-1.4
Experience in a related occupation	6.4	6.3	6.6	6.7	6.7	6.6	7.0	0.6
Long-term OJT (more than 12 months)	8.9	8.8	8.6	8.4	8.5	8.5	8.4	-0.5
Moderate-term OJT (1-12 months)	17.3	16.9	16.8	16.1	15.6	15.5	14.8	-2.5
Short-term OJT (less than 1 month)	41.4	41.3	41.0	41.1	41.7	41.8	42.4	1.0

Table 4. Employment and total job openings, 1998-2008, and 1997 median hourly earnings by education and training category

Education and training category	Total job openings* due to growth and replacement, 1998-2008		Percent distribution of median hourly earnings, 1997**			
	Number	Percent distribution	1	2	3	4
Total, all occupations	55,008	100.0	25.0	25.0	25.0	25.0
First professional degree	617	1.1	92.2	7.8	...	...
Doctoral degree	502	0.9	100.0	...	...	...
Master's degree	374	0.7	97.5	2.5	...	...
Work experience plus bachelor's or higher degree	3,372	6.1	94.1	3.2	2.7	...
Bachelor's degree	7,822	14.2	76.2	19.1	3.3	1.4
Associate degree	2,422	4.4	70.5	25.3	4.2	...
Postsecondary vocational training	1,680	3.1	7.2	60.5	17.2	15.1
Work experience in a related occupation	3,699	6.7	26.1	50.7	23.1	0.1
Long-term on-the-job training	4,411	8	15.9	57.7	7.3	19.1
Moderate-term on-the-job training	6,218	11.3	0.8	55.9	39.8	3.6
Short-term on-the-job training	23,890	43.4	0.7	7.8	35.8	55.8

*Total job openings represent the sum of employment increases and net replacements. If employment change is negative, job openings due to growth are zero and total job openings equal net replacements.

** The quartile rankings of Occupational Employment Statistics hourly earnings data are presented in the following categories: 1=very high ($16.25 and over), 2=high ($10.89 to $16.14), 3=low ($7.78 to $10.88), and 4=very low (up to $7.76).The rankings are based on quartiles using one-fourth of total employment to define each quartile.

Table 5. Weighted average percent change in skill requirements by sector, 1988-1998

Sector	*1988-1992*	*1992-1996*	*1996-1998*	*1988-1998*
All private nonagricultural industries	0.2	0.3	0.2	0.7
Goods producing	1.1	-0.1	-0.2	0.7
Mining and construction	0.6	0.2	-1.6	-0.8
Manufacturing	1.2	-0.2	0.3	1.2
Services producing	-0.1	0.4	0.4	0.8
Transportation, communications and utilities	0.9	0.8	1.5	3.1
Trade	-0.3	-0.1	-0.2	-0.5
Finance, insurance and real estate	0.0	0.4	1.3	1.7
Services	-0.1	0.7	0.7	1.3

Table 6. Economy-wide change in skill requirements, 1988-1998

	1988-1992	1992-1996	1996-1998	1988-1998
Total	0.5	0.1	0.5	1.0
Intra-industry	0.2	0.3	0.2	0.7
Interindustry	0.3	-0.2	0.3	0.3

MACROECONOMIC AND REGIONAL FORECASTING ISSUES

Chair: Paul Sundell
Economic Research Service, U.S. Department of Agriculture

Do Region-Specific Exchange Rate Indices Improve Regional Forecasts? The Case of State-Level Manufacturing Employment,
Amanda Hollenbacher, Lycoming College
Azure Reaser, Bureau of Labor Statistics, U.S. Department of Labor
David B. Yerger, Lycoming College

Are Rising Farm Prices Useful Inflation Indicators: the 1970s, 1980s, and 1990s?,
David Torgerson, Economic Research Service, Department of Agriculture

An Improved Phase Plane Model of the Business Cycle,
Foster Morrison and Nancy L. Morrison, Turtle Hollow Associates, Inc.

DO REGION-SPECIFIC EXCHANGE RATE INDICES IMPROVE REGIONAL FORECASTS? THE CASE OF STATE-LEVEL MANUFACTURING EMPLOYMENT

Amanda Hollenbacher, Lycoming College
Azure Reaser, Bureau of Labor Statistics
David B. Yerger, Lycoming College

I INTRODUCTION

This paper analyzes the impact of exchange rate movements upon state level manufacturing employment over a 25-year period ending in 1998. The model upon which the estimation is based was developed in the mid 1980's by Branson and Love (1986). This study extends their work in three ways. First, by extending the sample period beyond their 1986:1 ending point, our sample captures both sides of the 1980's world oil price spike and the U.S. dollar spike of the 1980's. This reduces the odds of spurious correlation being responsible for any findings of adverse impacts from exchange rate or energy price movements. A second extension of the work is that the model is estimated at the state level separately for durable goods and non-durable goods manufacturing, and not just for all manufacturing employment as in Branson and Love (1987).

The final extension of this paper is to estimate the model using both a national exchange rate index and region-specific exchange rate indices based on the work of Hervey and Strauss (1998a). (Our thanks to Hervey and Strauss for providing us with their exchange rate data.) Several recent papers have shown that export weighted region-specific exchange rate indices within the U.S. differ in their pattern of movements from a national exchange rate index. To date, however, very little work exists in the literature investigating whether these region specific exchange rate variables improve the fit or forecasting ability of regional economic models. This paper is one of the first, of which the authors are aware, to directly test for improved explanatory power from regional economic models utilizing regional rather than national exchange rate measures.

The relevant literature is briefly reviewed in the next section of the paper and the model itself, and the data used, is outlined in section III. Section IV contains a summary of the key empirical results and section V concludes the paper.

II LITERATURE REVIEW

The impact of the sharp spike during the first half of the 1980's in the value of the U.S. dollar upon U.S. employment in trade sensitive sectors was the focus of much investigation in subsequent years. A small literature examined the impact of the dollar movements at the state or regional level. In a series of papers, Branson and Love (1986, 1987) tested the impact of the dollar movements on U.S. manufacturing employment at either the state, or industry-specific level. They derived a reduced form model of manufacturing employment as a function of business cycle variables, the real price of energy, and the national real exchange rate. When estimating the model with quarterly data from 1970:1 to 1986:1 for all manufacturing employment at the state level, they find the elasticity of employment with respect to an appreciating dollar to be negative and statistically significant in 36 of 51 cases (all states + D.C.). Based on their parameter estimates, they find that the dollar's appreciation from 1980 to 1985 lead to a loss of approximately one million manufacturing jobs over this period.

Carlino (1990) estimated the impact of exchange rate movements upon the growth rates of Gross State Product (GSP) using annual data over the 1973-86 period. GSP growth rates were estimated as a function of U.S. and foreign real GDP growth rates, U.S. and foreign labor productivity growth rates, and a national real exchange rate index. In contrast to Branson and Love's findings of widespread negative effects on manufacturing employment, Carlino finds an adverse effect on GSP growth rates in only seven states. A positive effect is found in four states. The reduced frequency of adverse effects at the state level is not surprising given the many non traded goods sectors that are part of the GSP computation relative to the manufacturing sector.

In the 1990's a different strand of literature developed in which various U.S. region-specific real exchange rate measures were constructed and their movements contrasted against one another as well as traditional national exchange rate measures. These studies have established that the regional indices at least in part move independently of one another and the national index.

Clark, Sawyer, and Sprinkle (1997) constructed a quarterly export-weighted real exchange rate measure from 1973:3 to 1994:4 for the 'Southern Dollar' based on state level exports. The Southern Dollar region included all states that are former members of the Confederate States of America. They found that the Southern Dollar real exchange rate index and an index comprising the rest of the U.S. are not cointegrated. Moreover, the rest of U.S. index is not causing movements in the Southern Dollar index in the Granger-causality sense. Clark, Sawyer, and Sprinkle (1999) then extend the study by computing export-weighted real exchange rate indices for each of the nine census regions over the same period. They find that the national index is cointegrated with only two of the nine regional indices and that the national index is Granger-causing movements in only one of the nine regional indices.

Hervey and Strauss (1998a) construct export-weighted real exchange rate indices for the eight BEA regions and the entire U.S. using monthly data from 1970.1 to 1996.12. For each geographic unit, three indices are created based on the region's exports of all manufacturing goods, durable goods only, or non-durable goods only. They find that significant differences exist in the pattern of movements in the regional indices. In particular, the Midwest and Southwest region have faced an appreciating trend in their dollar since 1974. These two regions did not see the same type of decline in the value of their dollar post 1985 as did the other regions and the entire U.S. Hence, the stabilization (recovery?) of the manufacturing sector in the Midwest in the 1990's cannot be attributed to improvements in the region's real exchange rate.

In a follow-up paper Hervey and Strauss (1998b) use these region-specific exchange rate measures to test for the impact of changes in real exchange rates and foreign incomes upon regional manufacturing output for the eight BEA regions with annual data from 1970-1997. They estimate the impact upon four different measures of regional output: total gross regional product (GRP), GRP attributable to manufacturing, GRP attributable to durable goods manufacturing, and GRP attributable to non-durable goods manufacturing. They find minimal evidence of an impact from real exchange rate movements. Of the 32 region estimates, a negative effect from exchange rate movements was found in only three cases (Mideast durable GRP, Southwest manufacturing GRP and non durable GRP) while a positive effect was found in five cases (New England manufacturing GRP and non durable GRP, Mideast non durable GRP, Southeast manufacturing GRP and durable GRP).

Prior to this study the only work, of which the authors are aware, that directly compared the performance of an economic model using both national and region-specific real exchange rates was by Cronovich and Gazel (1998). They first create region specific export-weighted annual real exchange rates for the 50 states and D.C. over the 1987-1991 period. They then estimate a fixed effects panel model of state manufacturing exports as a function of: gross state product, state-specific real exchange rates, and state-specific measures of foreign income in export markets. If a national exchange rate measure is used, the exchange rate is not a significant determinate of state manufacturing exports. When the state-specific exchange rate measures are used, however, then a dollar appreciation has a negative and significant impact upon state level exports. Out of sample forecasting using state-specific exchange rates also was found to be superior to forecasts using a national exchange rate index on the basis of smaller out of sample forecast errors.

While Cronovich and Gazel's findings do show improved model performance from the use of region-specific exchange rate measures, the generalizations that can be drawn from their study are limited. They focus upon that slice of economic activity most likely to be impacted by currency movements, manufacturing exports, and estimate their model over only a few years of data.

This study will examine if Cronovich and Gazel's finding of the superiority of region-specific exchange rate indices continues to hold if a much longer time period is analyzed, 1974-1998, and if manufacturing employment, rather than exports, is the dependent variable. The region-specific real exchange rate variables from Hervey and Strauss will be used in the reduced form model of Branson and Love. The fit and forecasting ability of the model will be compared using both national exchange rate indices and the region-specific indices to see if meaningful differences exist.

III MODEL

As noted previously, the model is taken directly from Branson and Love (1986). For a complete derivation of the model, see their paper. Export supply is specified as a function of the real wage while export demand is a function of relative home versus foreign prices. Foreign income and real interest rate variables were dropped from the model as they added no explanatory power, and did not meaningfully change the estimated elasticities of employment with respect to exchange rate movements.

The estimated reduced form version of the model is given below as equation one. The dependent variable is the natural logarithm of employment. The explanatory variables include a constant, the real exchange rate, and three variables to capture secular, cyclical, and potential structural changes in demand. A trend term captures secular changes, the log of the unemployment rate is used to capture business cycle effects, and the real price of energy is included to capture the impact of major factor price shocks.

$$(1) \qquad y_{it} = \beta_0 + \beta_1 t + \sum_{j=0}^{4} \beta_{2j} LURT_{t-j} + \sum_{k=0}^{4} \beta_{3k} LRENGY_{t-k} + \sum_{l=0}^{6} \beta_{4l} LREX_{t-l} + \mu_t$$

where:
y_{it} =log of employment in state i,
t = trend variable
LURT = log of the national unemployment rate
LRENGY = log of the national real price of energy
LREX = log of the real exchange rate index
μ_t = *the error term*

The data is quarterly from 1974:1 to 1998:4. The employment data is the number of employed workers and is from the Bureau of Labor Statistics' *Employment and Earnings*. Three different measures of employment are used: total manufacturing, durable goods only, and non-durable goods only. The real energy index is the PPI for energy divided by the CPI-Urban index for all consumer goods.

The real exchange rate measures are from *Hervey and Strauss (1998a). The model is estimated* first using a national exchange rate index and then estimated again using the region-specific exchange rate measure. Note that when the dependent variable is either durable or non-durable goods employment that the exchange rate measures are based solely on exports of those goods.

In sum, a total of 306 different versions of equation (1) are estimated given the 51 states/D.C., three different employment measures, and two different real exchange rate measures. All models are estimated *using an AR(1) correction as in Branson and Love* (1987). Original estimates from OLS indicated significant serial correlation problems. For the all manufacturing with a national exchange rate index case, the null hypothesis of no serial correlation was rejected in 49 of 51 estimates. After estimating the model with the AR(1) correction the null of no serial correlation is never rejected.

The key results from these estimations are reported in the next section. Space constraints preclude presenting the complete econometric estimation results for each equation, but these results are available from the authors upon request.

IV RESULTS

Comparison with Branson & Love

Equation (1) initially is estimated using all manufacturing employment as the dependent variable and a national exchange rate index as this version is the most direct extension of Branson and Love's work. The results are summarized and contrasted with B&L's findings in Table 1. The signs of the trend terms are consistent with the well-known decline in the traditional manufacturing region of the U.S. and the effect is even more pronounced in our study than in B&L. While B&L found a negative and significant trend in just 11 states, the trend was negative in 27 states in our study. Moreover, the negative terms were concentrated in the New England, Mid East, and Great Lakes regions where 16 of the 17 states had negative trends. The parameter estimates on the unemployment variable were as expected, nearly always negative and statistically significant, with few differences between the two studies. Extending the sample to include both sides of the early 1980's oil price spike, however, eliminated any findings of an adverse impact from energy prices in our study whereas B&L had found a negative impact on several states.

The negative impact from exchange rate movements also was less frequent in this study than in B&L. Capturing both sides of the dollar's 1980's spike reduced the findings of an adverse effect on employment from 36 states in B&L to 27 states in this study. Also, this study finds a positive effect from a dollar appreciation in 12 states, six of which are in the New England and Mid East regions, while B&L found a positive effect in only one state. Overall, this study still finds fairly widespread adverse effects on state level manufacturing employment from an appreciating dollar, albeit at a diminished level relative to B&L.

Do Region-Specific FX Rates Improve Fit?

There is minimal evidence that estimating equation (1) for the all manufacturing employment case using region-specific values for LREX rather than the national index improves the fit of the model. The results are summarized in Table 2. In 29 of the 51 states/D.C., the adjusted R^2 is higher using the region-specific exchange rate than the adjusted R^2 when the national exchange rate index is used. A rank sign test of the null hypothesis that there is no difference in the adjusted R^2 of the region specific versus national exchange rate versions of the model across the states fails to reject the null as the p-value of the test statistic is 0.33. Nor is there any meaningful difference in the frequency of statistically significant parameter estimates across the two exchange rate versions of the model.

Estimating the model for durable goods employment only, or non-durable goods employment only, does not improve the performance of the regional exchange rates version of the model relative to the national exchange rate index version. (Recall, that the exchange rate indices are weighted by the exports of just durable goods or non-durable goods when constructing their respective indices.) As seen in Table 2, when non-durable goods employment is the dependent variable the adjusted R^2 is higher using the region-specific exchange rate in just 13 of 51 states. The only meaningful change in the frequency of statistically significant parameter estimates is that the prevalence of adverse effects on employment from a dollar appreciation declines from 35 of 51 to 28 of 51 cases when the region-specific exchange rates are used.

A similar drop in the frequency of adverse employment effects from a rising dollar when region-specific exchange rates are used is found for the durable goods employment estimates. The number of states with negative exchange rate coefficients declines from 19 using national exchange rate measures to only 8 using regional exchange rate measures. Nor is there any gain in the overall fit of the model when regional exchange rates are utilized. The adjusted R^2 is higher using the region-specific exchange rate in 24 of 51 states.

<u>Do Region-Specific FX Rates Improve Employment Forecasts?</u>

Forecasts for each of the 306 estimating equations were created by first estimating the model over the 1974:1 to 1994:4 period and then using the resultant parameter estimates to forecast manufacturing employment over the 1995:1 to 1998:4 quarters. The forecast performance of the region-specific exchange rates versus national exchange rate versions of the model are compared using Theil's Inequality Coefficient, U, which is computed as shown in equation (2).

$$(2) \qquad U = \{[(1/T)*\Sigma^T_{t=1}(Y^f_t - Y^a_t)^2]^{1/2}\} / \{[(1/T)*\Sigma^T_{t=1}(Y^f_t)^2]^{1/2} + [(1/T)*\Sigma^T_{t=1}(Y^a_t)^2]^{1/2}\}$$

Y^f_t is the forecasted value at period t and Y^a_t is the actual value at period t. Note that the numerator of U is simply the root mean squared error of the forecast. Theil (1961) shows that U is bounded by 0 and 1 with a 0 indicating a perfect fit between the forecasted and actual values.

The forecast comparisons are summarized in Table 3 which shows the value of U for each possible national exchange rate versus region-specific exchange rate pairing. There is no evidence that the use of region-specific exchange rate variables improves the model's employment forecasting ability. When all manufacturing employment is the dependent variable, U is lower for the region-specific exchange rate version of the model in 26 of 51 states. When the dependent variable is either durable goods employment only or non-durable goods only, the case for region-specific exchange rates is even weaker. U is lower using regional exchange rates in 20 of 51 states for durable goods employment, but in only 11 of 51 cases for non-durable goods employment.

V CONCLUSION

This study updates literature from the latter 1980's on the impact of exchange rate movements upon U.S. manufacturing employment. It finds that once the data set is extended to include data beyond the peak value of the dollar in the middle 1980's, the prevalence of adverse effects from an appreciating dollar declines. Branson and Love's data ended in the first quarter of 1985, and they found a negative effect on state level manufacturing employment from an appreciating dollar in 36 of 51 states/D.C. This study extends the sample through the fourth quarter of 1998 and finds adverse exchange rate effects for 27 of 51 states/D.C.

The prevalence of adverse exchange rate effects weakens further if one focuses upon durable goods manufacturing employment rather than total manufacturing employment. When a national exchange rate measure is used, adverse employment effects are found in only 19 states. When region-specific exchange rate measures are used, adverse effects decline further to just 8 of 51 states/D.C. While a strong appreciation of the dollar would have negative effects upon manufacturing employment in a number of states, the extent of the employment decline is likely to be less widespread than suggested by earlier analyses.

The second primary objective of this study was to test if the utilization of region-specific exchange rate variables, rather than a national exchange rate index, improved the fit or forecasting ability of the model. Very little evidence was found to support claims of superior model performance when region-specific exchange rate measures were used. Forecasts of state level manufacturing employment- whether total, durable goods only, or non-durable goods only- simply were not meaningfully improved by using region-specific exchange rates. These results contrast with the findings of Cronovich and Gazel (1998), but this study differed from theirs in at least two important respects. First, their time period was much shorter covering just 1987-91. Also, their dependent variable was state manufacturing exports, not employment. With exports as the dependent variable it is more likely one would find an impact from exchange rate movements.

In sum, while regional exchange rates may differ in their movements from one another and from a national index, this paper's findings question whether these differences are large enough to meaningfully improve the accuracy of most models of regional economic activity.

REFERENCES

Branson, William H. and James P. Love. 1987. "The Real Exchange Rate and Employment in U.S. Manufacturing: State and Regional Results," *NBER Working Paper # 2435*. National Bureau of Economic Research. Cambridge, MA.

Branson, William H. and James P. Love. 1986. "Dollar Appreciation and Manufacturing Employment and Output," *NBER Working Paper # 1972*. National Bureau of Economic Research. Cambridge, MA.

Carlino, Gerald A. 1990. "Should States Fear the Effects of a Changing Dollar?" *Business Review*. Federal Reserve Bank of Philadelphia. Sept/Oct. pp. 3-12.

Clark, Don P., Charles Sawyer, and Richard L. Sprinkle. 1999. "Regional Exchange Rate Indexes For The United States," *Journal of Regional Science*. Vol. 39. No. 1. pp. 149-166.

Clark, Don P., Charles Sawyer, and Richard L. Sprinkle. 1997. "The Value of the Southern Dollar," *The Review of Regional Studies*. Vol. 27. No. 2. pp. 185-193

Cronovich, Ron and Ricardo Gazel. 1998. "Do Exchange Rates and Foreign Income Matter For Exports At The State Level?" *Journal of Regional Science*. Vol. 38. No. 4. pp. 639-657.

Hervey, Jack L. and William A. Strauss. 1998b. "Foreign Growth, the dollar, and regional economies, 1970-97," *Economic Perspectives*. Federal Reserve Bank of Chicago. Fourth Quarter. pp. 35-54.

Hervey, Jack L. and William A. Strauss. 1998a. "A regional export-weighted dollar: An examination of the regional impact of exchange rate change," *International Regional Science Review*. Vol. 21. No. 1. pp. 1-20.

Theil, H. 1961. *Economic Forecasts and Policy*. Amsterdam: North-Holland. pp. 30-37.

Table 1-- Comparison of Model Estimates With Branson & Love's Results
Estimates For All Manufacturing Employment
Using National Exchange Rate Index For LREX

Frequency of Statistically Significant Parameter Estimates

Variable	B&L	This Study	Variable	B&L	This Study
Trend			$\sum$ LERGY		
# insignif.	10	0	# insignif.	27	37
# - & signif.	11	27	# - & signif.	7	0
# + & signi	30	24	# + & signif.	17	14
$\sum$ LURT			$\sum$ LREX		
# insignif.	3	6	# insignif.	14	12
# - & signif.	48	45	# - & signif.	36	27
# + & signi	0	0	# + & signif.	1	12

Spatial Pattern of Trend Term Signs

- means negative significance at 10% level, + means positive significance at 10% level, blank means not significant at 10% level

Region/State	B&L	This Study	Region/State	B&L	This Study
New England			Southeast		
CT	-	-	AL	+	+
ME		-	AR	+	+
MA		-	FL	+	-
NH	+	-	GA	+	+
RI		-	KY	+	+
VT	+	-	LA	-	+
Mid East			MS	+	+
DE		-	NC	+	-
DC	-	-	SC		-
MD	-	-	TN	+	-
NJ	-	-	VA	+	-
NY	-	-	WV	-	-
PA	-	-	Southwest		
Great Lakes			AZ	+	+
IL	-	-	NM	+	-
IN		-	OK	+	+
MI	+	-	TX	+	+
OH	-	-	Mountain		
WI	+	+	CO	+	+
Plains			ID	+	+
IA		+	MT		-
KS	+	+	UT	+	+
MN	+	+	WY	-	+
MO	+	-	Far West		
NE	+	+	AK		+
ND	+	+	CA	+	-
SD	+	+	HA		-
			NV	+	+
			OR	+	+
			WA	+	+

Table 2-- Comparison Of Model Estimates Using National Versus Regional Exchange Rates

Variable	All Manufacturing Employment Nat'l FX	All Manufacturing Employment Reg'l FX	Non-Durable Goods Employment Nat'l FX	Non-Durable Goods Employment Reg'l FX	Durable Goods Employment Nat'l FX	Durable Goods Employment Reg'l FX
ΣLURT						
# insignif.	6	1	23	28	1	4
# - & signif.	45	50	26	22	50	47
# + & signif.	0	0	2	1	0	0
ΣLERGY						
# insignif.	37	31	37	42	29	28
# - & signif.	0	0	0	1	2	2
# + & signif.	14	20	14	8	20	21
ΣLREX						
# insignif.	12	15	7	16	26	24
# - & signif.	27	27	35	28	19	8
# + & signif.	12	9	9	7	6	19
# of states for which adj. R^2 using Reg'l FX rates is > than adj. R^2 when using Nat'l FX rates		29		13		24

Note: Significance is taken to be pvalue <= 0.10

Table 3-- Comparison Of Forecast Performance Using National Versus Regional Exchange Rates
Forecast Performance Measured Using Theil's Inequality Coefficient U
Reported Value is U

State	All Manufacturing Employment Nat'l FX	All Manufacturing Employment Reg'l FX	Non-Durable Goods Employment Nat'l FX	Non-Durable Goods Employment Reg'l FX	Durable Goods Employment Nat'l FX	Durable Goods Employment Reg'l FX
AL	.009	.010	.019	.021	.007	.008
AK	.085	.003	.657	.301	.337	.288
AZ	.003	.012	.022	.024	.006	.006
AR	.005	.010	.012	.013	.005	.006
CA	.006	.003	.008	.007	.014	.014
CO	.005	.003	.022	.022	.010	.010
CT	.006	.004	.003	.005	.014	.015
DE	.030	.005	.024	.026	.068	.067
DC	.005	.005	.126	.145	.471	.642
FL	.003	.010	.023	.023	.004	.003
GA	.004	.010	.008	.008	.005	.006
HA	.031	.003	.037	.042	.171	.173
ID	.003	.003	.041	.042	.028	.022
IL	.002	.003	.003	.003	.007	.006
IN	.003	.003	.010	.010	.004	.004
IA	.002	.002	.010	.012	.013	.009
KS	.004	.002	.002	.002	.013	.012
KY	.003	.011	.009	.011	.006	.004
LA	.002	.010	.003	.005	.009	.007
ME	.008	.004	.011	.007	.022	.021
MD	.005	.005	.004	.003	.005	.004
MA	.007	.004	.009	.007	.021	.020
MI	.006	.003	.006	.006	.009	.009
MN	.003	.002	.007	.006	.002	.003
MS	.017	.011	.031	.033	.018	.018
MO	.004	.002	.009	.009	.004	.005
MT	.021	.003	.040	.044	.031	.031
NE	.001	.002	.011	.010	.002	.002
NV	.008	.003	.009	.010	.015	.013
NH	.011	.004	.026	.025	.014	.012
NJ	.003	.005	.012	.012	.010	.010
NM	.016	.012	.017	.022	.030	.032
NY	.004	.005	.006	.006	.005	.005
NC	.009	.110	.014	.015	.007	.007
ND	.007	.002	.022	.021	.026	.028
OH	.002	.003	.001	.002	.004	.004
OK	.004	.012	.011	.010	.008	.006
OR	.007	.003	.012	.012	.013	.013

Table 3 Continued

State	All Manufacturing Employment Nat'l FX	All Manufacturing Employment Reg'l FX	Non-Durable Goods Employment Nat'l FX	Non-Durable Goods Employment Reg'l FX	Durable Goods Employment Nat'l FX	Durable Goods Employment Reg'l FX
PA	.004	.005	.003	.004	.008	.006
RI	.013	.004	.010	.011	.004	.005
SC	.009	.011	.016	.017	.008	.008
SD	.008	.002	.012	.012	.010	.009
TN	.010	.011	.018	.020	.009	.010
TX	.002	.012	.003	.003	.005	.009
UT	.004	.003	.007	.012	.007	.007
VT	.009	.004	.030	.030	.044	.027
VI	.007	.011	.011	.010	.009	.005
WA	.016	.003	.015	.016	.018	.021
WV	.008	.011	.007	.007	.011	.009
WI	.003	.003	.006	.006	.003	.003
WY	.024	.003	.076	.099	.024	.026

Are Rising Farm Prices Useful Inflation Indicators: the 1970s and 1980s and 1990s?

David Torgerson, Economic Research Service, USDA[1]

Overall Inflation and Farm Commodity Price Variations

The focus of this analysis is the role that farm commodity prices (wholesale prices for raw farm commodities such as grains, fruits, vegetables, tobacco and other raw materials grown on farm) have had in indicating future inflation. Hamilton (2000) and Hooker (1999) showed that the oil price GDP growth link was dramatically altered in 1983. Following Hamilton and Hooker, I ask if the inflation-farm price link changed in the 1980s and 1990s compared with the 1970s. Farm commodity inflation may reflect other forces that have a role in determining inflation. Farming accounts for about 2 percent of U.S. GDP. Indeed the entire food and fiber system, from farm to final consumer, is less than 20 percent of GDP. So, the channels for substantial transmission of raw farm price increases to overall inflation are limited.

Nevertheless, farm price inflation has sometimes been an early indicator of a build up of inflationary pressures. The record high farm prices in 1946 presaged a significant ratcheting up of inflation in the late 1940s. Indeed, commodity price inflation has been a useful predictor of overall inflation throughout much of U.S. history. As agriculture, oil extraction and other raw material mining have declined relative to the overall economy the link between commodity and overall inflation has apparently weakened. Yet large increases in commodity prices may indeed continue to influence overall inflation. The changing role of crude oil prices in determining U.S. economic growth has been well documented by Hamilton (2000). As recently as the early 1980s oil prices played a significant role in U. S. economic growth (Hooker (1999)). By some accounts high and rising real oil prices accounted for up to 45 percent of the GDP decline from the back-to-back recessions of the early 1980s. A Brookings Institution study done by Bosworth and Lawrence (1982) (henceforth Bosworth) demonstrated that industrial and farm commodity price inflation played a key role in the inflation of the 1980s.

[1] The author is an agricultural economist with the Economic Research Service.

Rising primary farm commodity prices may still be an indicator of future run-ups in industrial commodity prices. For example, if the dollar is greatly undervalued, U.S. farm prices may be low relative to other producers and potential exporters' prices. This situation would induce a rise in the demand for U.S. farm commodities that would drive up domestic farm prices. Other raw materials prices could be similarly lifted by a weak dollar thereby boosting consumer prices. Besides the direct pressure higher farm prices have on domestic inflation through higher priced food and fiber products, the impact of higher industrial raw material prices on other final products would show up later, further boosting overall inflation.

The forces connecting crude farm prices and inflation may have weakened over time, as farming has become a smaller part of the economy. First, the domestic food and fiber system takes a smaller portion of the consumer food dollar, making the potential for a direct impact of raw farm prices on inflation smaller than in the 1970s. Secondly, with the widespread deregulation in farming and transportation and liberalization of world trade the ability to pass through crude good price increases has diminished over time.

This work seeks to test if the 1970s links between crude farm prices and overall inflation, and farm price inflation and overall inflation were useful in forecasting aggregate inflation. (The measure of farm commodity prices was the crude farm Producer Price index--reflecting the wholesale price of farm-produced crops and livestock.) Further, do these links improve inflation forecasts in the 1980s and 1990s? I replicate Bosworth using the farm price PPI instead of the overall international commodity index used by Brookings. I then compare the above forecasting results to results obtained using only the macroeconomic variables used by Bosworth. This process generates the six models below summarized in equations (1) to (6), with alternative equations (1) and (2) estimated from 1948 to 1969, (3) and (4) estimated over 1958 to 1979, and (5) and (6) estimated over the 1968 to 1989 period.

Brookings' Model

The Bosworth model had the quarterly GDP deflator equation (P) dependent on:

(1) inflation from the previous quarter, P(-1),

(2) contemporaneous Aggregate Demand pressure measured as the prior quarter's GDP divided by trend GDP from the prior quarter,

(3) the change in Aggregate Demand pressure from the prior quarter's aggregate demand pressure and

(4) a contemporaneous two-year weighted average of the change of the United Nations world raw material price index,

(5) the United Nations raw material price index one period prior.

The use of a lagged dependent variable such as P (-1) may invalidate in-sample test statistics, so those significance tests reported in Bosworth are biased. Yet, the use of a lagged dependent variable will often improve forecasts of the dependent variable and is standard practice in building macroeconomic models. A macroeconomic variable in the current quarter tends to be related to that variable in the preceding quarter. Major forces take time to work themselves through the economy. While both industrial and farm commodity prices have a tendency to bounce around from quarter to quarter, overall inflation in one-quarter will typically help predict inflation in the next quarter. A commodity price is like a speedboat that is highly maneuverable, while aggregate inflation is like an ocean liner changing direction and speed relatively slowly.

The ratio of GDP relative to trend GDP measures the tightening of labor and input markets. Businesses raise output thereby bidding up wages. Since wages are the largest cost, businesses pass on the higher wage costs in higher prices. Similarly, bottlenecks in other major input costs such as rent and intermediate materials push up overall prices as the economy moves at or above full capacity. The current economic situation shows only a slight increase in inflation despite a high GDP to trend GDP ratio only because of extraordinary increases in productivity growth. Other more sophisticated measures of demand pressure, such as implied demand for capital in a manufacturing sector compared to capital stock in that sector, are usually not available until years after a forecast has to be made.

Again, a large change in aggregate demand quarter to quarter may steer the macroeconomic ship toward higher inflation if demand pressure is up sharply from the prior quarter. The adjustment costs accrued in moving to a higher level of output than the prior quarter will typically be passed on in higher inflation. On the other hand, the economy may slow if hit with too large an upward movement in aggregate demand from one quarter to the next putting downward pressure on inflation. Looking at the results for the countries analyzed in Bosworth, the picture for OECD is mixed with the United States and most of the developed economies seeming to have a positive sign for a one-quarter change in aggregate demand pressure. France and Japan, unlike the other developed economies, have negative signs on a one-quarter change in aggregate demand.

The commodity price inflation (and change in commodity price inflation) represents at least an early warning of higher input prices or increased economic tightness not yet reflected in aggregate demand. Generally, the impact on overall inflation of commodity prices compared to aggregate demand variables should be small.

The Brookings story says inflation has a life of its own, with aggregate demand pressure an important variable, with sharp changes in aggregate demand pressure boosting or slowing inflation and commodity price inflation and changing commodity price inflation modestly boosting overall inflation. Bosworth (1982) also demonstrates in several ways that commodity pricing and inflation were linked at least in the 1970s.

The Base Replication Estimated over 1948 to 1969, 1970s Forecast Comparison

I replicated the Brookings approach by using the Macroeconomic variables used in Bosworth and the PPI farm price index. I developed a model that met the following criteria estimated over 1949I to 1969IV (the first quarter of 1949 to the fourth quarter of 1969):

Select the equation that minimizes the Akaike Information Criteria (AIC)[2] subject to:

a. No theoretically inappropriate signs allowed

b. Every equation has a constant term no matter what its T-statistic is

c. P(-1) is forced to enter every equation; and

d. Each estimated T-statistic must have an estimated probability of a false positive below 20 percent.

The lowest AIC (as measured by EViews) equation fitting the four criteria was then selected by iterating over 6 quarters. I refer to this estimated equation as MACROAG4869 reflecting the inclusion of data for the 1948 to 1969 including the macro variables of the Brookings study (P up to 6 lags, (GDP/trend GDP) up to 6 lags and the change in (GDP/trend GDP) up to 6 lags). Instead of the aggregate commodity index and its inflation, farm PPI and farm PPI inflation were used since the focus is specifically on the farm price and overall inflation link. (The UN commodity index also has oil and other mineral prices as well as other raw materials.)

Criteria a, b, and c. have been shown to improve out-of-sample forecasts in various Monte Carlo studies. I was forced to employ restricted lag length due to degree of freedom problems that would arise if longer lags were allowed. Further, other studies suggest most supply and demand shocks take at most 6 quarters to work through the economy (Abel and Bernanke (1999)).

MACROAG4869, the lowest AIC equation for GDP deflator inflation given the data and variables above and maximum lag length of 6, was:

[2] The AIC used here is as reported by the EViews computer package. The EViews measure of the AIC is the negative natural log of the AIC shown in most textbooks such as Diebold (1998). EViews is a licensed econometric package available from Quantitative Micro Systems.

$$(1)\quad P = .3504 * P(-1) + .0653 * AD(-1) + .0003 * PPIFARM + .5733$$

where P is the quarterly GDP deflator, AD(-1) is the ratio of GDP in the prior quarter to trend GDP in the prior quarter, and PPIFARM is the current quarter producer price received by farmers. Although *PPIFARMINF, the* inflation in the PPIFARM price index and 6 lags were tested none had a lower AIC, fitting the four side criteria, than the selected equation (1).

The model was then used to forecast the 1970s. While the model is not re-estimated, history is updated to avoid cumulative errors. This procedure reflects the information set available to an analyst at the end of quarter being forecasted. (This is a more stringent than the Bosworth which included contemporary exogenous variables which are typically unavailable until the next quarter.)

Stock, bond, and commodity markets indicate the actual GDP inflation estimate release is a variable that moves markets. So this equation is of some significance to applied forecasters not just giving insight into happenings of the 1970s.

Now the same process was done excluding the PPIFARM and inflation in PPIFARM otherwise including the same potential variables as in (1)

$$(2)\quad P = .4744 * P(-1) + .0466 * AD(-1) + .1002 * CHAD(-1) + .4820$$

where variables are as defined above with CHAD(-1), the difference between
AD(-1) and AD(-2). We refer to this model as MACRO4869, the lowest AIC of all the models including up to 6 lags of P, AD, and CHAD.

As is shown in table 1, the MACROAG4869 simulated for the 1970s out-performed MACRO4869 with an out of sample error over the 1970s period that was one-third smaller. However, for the other two forecast periods

MACRO4869 was the superior model with root mean squared errors less than half and less than 10 percent the size of the MACROAG4869 simulations for the 1980s and 1990s respectively.

Re-estimated Models Estimated over 1958 to 1979, 1980s Forecast Comparison

As is standard practice, macroeconomic models are re-estimated incorporating ten new years and discarding the ten oldest years in the sample. So models are estimated for the 1958I to 1979IV using the same schema as above. MACAG5879 is the lowest AIC model subject to the other criteria using the same variables as above estimated over the 1958I to 1979IV period. The resulting model is:

$$(3)\ P = .6755 * P(-1) + .0528 * AD(-1) + .0001 * PPIFARM + .0168 * PPIFARMINF + .2523$$

where PPIFARMINF is farm commodity inflation as measured by percent change in PPIFARM and other variables are defined as above. Note that both the level of the farm price index and the inflation in the farm price enter in comparison to the level of farm prices only in equation (1).

The competing model estimated over the same time frame denoted MAC5879 is

$$(4)\ P = .8868 * P(-1) + .0482 * AD(-2) + .0515 * CHAD(-1) + .0663$$

where variables are defined above with AD(-2) as AD two quarters lagged.

The forecast competition for the 1970s and 1980s is won by MAC5879, as its root mean squared error is less than a quarter of that for MACAG5879 (table1). Indeed, MAC5879 is superior in forecasting to either of the models estimated over the 1948 to 1968.

Re-estimated Models over 1968 to 1989, 1990s Forecast Comparison

Again selecting the best model with the data from 1968 to 1989 including all variables results in

(5) $P = .8103 * P(-1) + .0219 * AD(-2) + .0133 * PPIFARMINF + .0236 * PPIFARMINF(-4) + .0122$.

We denote this model MACAG6889

The corresponding model MAC6889 is:

(6) $P = .8242 * P(-1) + .0424 * AD(-2) + .1358$.

MACROAG6889 out-forecasted MACRO6889 for the 1990s with a 15 percent lower Root mean squared error.

Best models for the 1970s, 1980s and 1990s and Interpretation

Table 1 allows comparison across models and time periods. The best model in forecasting out of sample for the 1970s was MACROAG4869, reflecting the importance of farm commodity prices in forecasting the inflation of the 1970s. For the 1980s MACRO5879 produced the best forecasts largely as expected.

Surprisingly, the best forecasting model for the 1990s was MACRO4869. Adding the noise of the data of the 1970s and 1980s apparently made the model deteriorate significantly. It could well be the extreme turbulence of oil and commodity prices in the 1970s and the overvalued dollar and tight monetary policy with loose fiscal policy and world debt crisis of the 1980s induced abnormal relations between commodity prices and inflation that was best to ignore. The root mean squared error of MACRO4869 was a mere 43 percent of the next best alternative MACRO5879. So while there is out-of-sample gain in dropping the data from the 1980s there is more to be gained from dropping the 1970s as well.

These qualitative results are unchanged if one re-estimates these models adding ten years and twenty years respectively so the estimation periods are 1948 to 1969, 1948 to 1979 and 1948 to 1989 instead of following the usual convention of dropping the oldest ten years when ten new years are available. Results are available from the author upon request.

Using farm prices provided superior forecasting for the 1970s, indicating that the sharp run up in inflation stating in the late 1940s that started with farm commodities gave a foretaste of the 1970s. So, indeed including a farm price index allowed superior inflation forecasting. For other periods, the 1980s and 1990s, using farm commodity prices in an inflation forecasting equation made for inferior forecasting performance. This is broadly consistent with Hamilton's and Hooker's findings that the influence of oil prices on productivity growth has become far less important than it was in the 1970 to 1982 period.

The results are similar to those discussing stock market returns where over some periods of time various strategies and types of funds outperform the market for a period of time only to be later beaten by the market. Forecasting inflation for longer periods of time shows the same problems as stock market forecasting. The use of farm prices helps forecasting inflation for the 1970s and 1990s if standard procedures are used. If one uses reasonable historical judgment and throws out more recent but less relevant data, the usual macroeconomic variables provide a superior base for forecasting in the 1980s and 1990s.

Unfortunately, this provides little guidance on the best model for the 2000s. Chechetti et al (2000) also shows how difficult it is to forecast inflation a year ahead. This is broadly consistent with the results here. The bottom line is the forecaster has to go beyond the past and extract the most relevant features of the past to forecast the future. The most difficult part of forecasting is fitting the stream of data that is evolving into the best framework. Some are better stock pickers for a period of time and some are better inflation forecasters but it is hard to maintain consistent superiority in either.

References

Abel, Andrew B. and Ben S. Bernanke. Macroeconomics , third edition, Addison Wesley 1998.

Bosworth, Barry P., and Robert Z. Lawrence. Commodity Prices and the New Inflation, The Brookings Institution 1982.

Chechetti, Stephan G., Rita S. Chu and Charles Steindel. "The Unpredictability of Inflation Indicators" Current Issues in Economics and Finance, Volume 6, no. 4, New York Federal Reserve Board Apr 2000.

Diebold, Francis X. Elements of Forecasting, SouthWestern College Publishing, 2000.

Hamilton, James D. "What is an Oil Shock", National Bureau of Economic Research Working Paper 7755, June 2000.

Hooker, Mark A. "Oil and the Macroeconomy Revisited", Federal Reserve Board International Finance Working Paper, Aug 1999.

Table1 OUT OF SAMPLE FORECASTING PERFORMANCE COMPARISON

	RMSE1970s	RMSE1980s	RMSE1990s
MACRO4869	0.011275	0.009208	**0.002281**
MACROAG4869	**0.007139**	0.023872	0.034856
MACRO5879		**0.003681**	0.005318
MACROAG5879		0.015248	0.023573
MACRO6889			0.007507
MACROAG6889			0.006381

RMSE19aas is root mean squared error computed from the first quarter to last quarter of the 19aas, aa=70,80,90
MACROiijj and MACROAGiijj estimated over the first quarter of 19ii to the last quarter of jj
BEST forecasting estimates for each decade in bold

Appendix Data Sources and Definitions

GDP is the real gross domestic product (a broad measure of the value of goods and services produced in the United States adjusted for overall price inflation) base year 1996 from the Bureau of Economic Analysis.

Trend GDP is the Hodrick-Prescott filtered GDP, computed by the Economic Research Service. Series available from author on request.

P is the GDP price deflator (the broadest measure of inflation in the U.S. economy) with the 1996 base year from the Bureau of Economic Analysis.

PPIFARM is the producer price index farm price 1982 base year from the Bureau of Labor Statistics.

PPIFARMINF is the percentage change of PPIFARM from prior quarter.

Data are the August 2000 releases downloaded from Haver Analytics.

AN IMPROVED PHASE PLANE MODEL OF THE BUSINESS CYCLE

Foster Morrison and Nancy L. Morrison
Turtle Hollow Associates, Inc.
PO Box 3639
Gaithersburg, MD 20885-3639 USA
Phone: 301-762-5652 Fax: 301-762-2044
Email: 71054.1061@compuserve.com

1. What Is the Business Cycle?

The systematic collection of econometric data began in the USA only after World War II. There are records of stock market prices that go back more than 100 years. Some commodity prices can be traced back several centuries. But anecdotal accounts of the business cycle can be found even earlier, such as the biblical story of Joseph in Egypt. The fact that the level of economic activity does not remain constant or just grow at a constant rate is a longstanding observation.

Significant fluctuations in economic activity create problems for both businesses and governments. When sales and tax revenues decline, operating expenses may fall only slightly or not at all. It may become necessary for organizations to either borrow money or curtail operations. Reducing expenses has the unfortunate effect of shrinking economic activity even more; it is a positive feedback.

The inconveniences and hardships created by the business cycle have generated an ongoing debate among economists and political theorists. Almost every conceivable action by governments, central banks, and businesses, including doing nothing, has its partisans. Most of these have been tried, at least in some watered down way, by various nations at various times. The net result has been that the business cycle usually responds, but it does not go away.

One basic fact is that the exchange economy is extremely dynamic. The physical sciences have outgrown the concept of a deterministic, "clockwork" universe, due to the success of quantum mechanics early in the 20th century and recent discoveries about chaotic dynamics (Gleick, 1987). Time series analysis is often used in economics, but the dynamics involved is obscured by the statistics. The dynamical interpretation of time series methods is noise-driven linear difference or differential equations (Jordan, 1972; Morrison, 1991).

If you want a simple, mechanical analogy for the economy, consider a system of belts and pulleys rather than clockwork. The belts stretch and slip on the pulleys, so the mechanism does not retain the rigid phase-locks of a gear train. Any would-be regulator wants to keep the belts adjusted to the optimum tension, but numerous individuals and organizations are tampering with the mechanism and inadvertently sabotaging the efforts.

Any attempts to ameliorate the business cycle should begin with some knowledge of its dynamics. Economic and political theorists have looked for simple control strategies, such as manipulating the money supply. But even for some fairly simple mechanical systems, optimal control can be counterintuitive. Recall that the way to pull a car out of a skid is to turn the wheels in the direction of the skid.

One practical consideration is that some industries are more cyclical than others. The same is true of government agencies. Many agencies are not affected in any way by moderate swings in the business cycle. Those running "safety net" operations may see their work loads climb when the business cycle dips. The Treasury Department, of course, is the executive agency most concerned with macroeconomic variation. Other agencies, such as Commerce and Labor, may collect the numbers, but what those numbers are does not affect their internal operations. The most significant participant in active macroeconomic management is the Federal Reserve System.

2. Capturing a Picture of the Business Cycle

The economy is a huge dynamic system with an unknown and probably unknowable number of variables. This is one reason that the collection of data is a recent phenomenon, even though exchange economies began in prehistoric times. Only with the invention of money did it become possible to measure all transactions on a common scale. And the daily variations of exchange rates, published in most newspapers, show that this scale is not as stable as the standard meter.

There are enormous difficulties to be surmounted in collecting econometric data. First of all, there is the difficulty of identifying something that can be measured. And then there is the effort required to do the measuring. For various reasons, businesses and individuals are often reluctant to provide information. In many cases the working economists have to be satisfied with incomplete data and must make extrapolations. A few types of data, such as stock market indices and commodity prices, are precise and easy to collect.

Exponential growth is the dominant dynamical characteristic of the U.S. and many other national economies. Some of this "growth" is due to persistent inflation, so the U.S. Department of Commerce issues inflation adjusted as well as current dollar estimates for the GDP (gross domestic product). The larger question of how long real exponential growth can continue has precipitated heated debates at times.

Nobody knows what negative feedback or combination of feedbacks will end economic growth. A combination of market forces, technological improvements, and government regulations has permitted growth to continue longer than some ecologists had expected. But there is always another feedback ready to come into play and the consequences of water shortages, especially in the western USA and other arid areas, are not yet known. Climate change, specifically global warming, is another topic stimulating extensive research and generating intense debates.

Prolonged economic growth and a booming stock market have decreased popular interest in the business cycle during the past few years. Some analysts believe that the business cycle has been smoothed out, claiming that the Federal Reserve finally has mastered the art of creating money at just the right pace. To address this hypothesis, however, it is necessary to have a qualitative model of the business cycle.

To supplement the quarterly releases of GDP numbers, the U.S. Department of Commerce introduced three composite indices of leading, coincident, and lagging indicators. The coincident index is a stand-in for the GDP, but it is normalized to average 100 over a predetermined period rather than being set to match the GDP. The other two indices are treated in the same way (*Handbook*, 1984). Taken together, these indices form a much simplified, three-dimensional model of the U.S. economy.

Each index has been constructed from a small number (21 currently) of carefully selected econometric series, just a few of the many available. These 21 (10 leading, 4 roughly coincident, and 7 lagging) are then reduced to just three numbers. The three indices have close to optimal reliability and signal-to-noise ratios. Decades of effort have been expended on constructing and maintaining these indices. Constructing a graphical phase plane plot of the cycle is a value added product that makes the indices easier to interpret.

Plotting the three indices as functions of time provides a useful tool for determining the state of the economy. But such representation is not optimal, either for detailed analysis or visual perception. A three-dimensional trajectory in 3-space, which could be created by computer plots of a stereo view, would yield a soaring, yet ragged helix.

Using logarithms of the data convert the helix to one of fairly even pitch (like the threads on a bolt). Data series exhibiting exponential growth, whether real, inflation created, or both, are best analyzed as logarithms. This converts the soaring arc of the exponential function into a straight line.

Trend removal collapses the helix into a hoop. The hoop is still three-dimensional, but it can then be projected onto an optimally oriented plane (or other surface), producing an easy to understand plot of the business cycle. Using logarithms of the data and doing the trend removal also yield results that can be analyzed by time series methods. The final two-dimensional phase plane plot produces a visual product that can be readily comprehended by users without extensive training in either economics or mathematics.

3. Specialized Tools and Techniques

The trend model used in developing this business cycle model is the low-pass ramp filter (Morrison and Morrison, 1997). This is a weighted mean, similar to the moving average, but it has been designed so that the end point rather than the middle point is the proper time reference for the filtered data.

The ramp filter is essential for analyzing and forecasting the most recent data. Trial and error has shown that a 60-point ramp filter is suitable for analyzing and forecasting the three indices and this spans 5 years. Two and a half years is way beyond the possible range of precise forecasting for these data, so a moving average is not usable. This is true of other econometric data, so the ramp filter is recommended for any and all such series or their logarithms, where appropriate.

Sixty points is not an optimal number for all series, but the correlation distance of detrended data will always be much shorter than the ramp filter length. The *correlation distance* (or *time*, in the case of econometric data) is that for which the ACF (autocorrelation function) drops to $1/e$ (0.367879...); it is a good measure of the range of forecast precision. Forecast reliability is another question, however.

A low-pass filter does not amplify noise, which differencing, especially higher-order differencing, will do. And unlike the case of polynomial regressions, other than a straight line, the extrapolation of the trend is plausible. Low-pass filters are also better trend models than regressions because the addition of new data does not alter the trend model for the earlier data. Deviations from the trend (and the error estimates) are readily transformed into a forecast for the initial series (and corresponding error estimates).

The earlier versions of this model used only the leading and coincident indices (Morrison and Morrison, 1997, 1998, 2000). With only two variables, constructing a phase plane plot is easy and without any ambiguities. However, there is some neglected information in the lagging index and the challenge is to access it while retaining the simplicity of a phase plane plot.

Several approaches were considered, but a simple projection onto a plane was chosen because of simplicity and computational stability. But first it is necessary to construct the detrended data points. The three indices used are: x = index of leading indicators, y = index of coincident indicators, z = index of lagging indicators. These are detrended and converted to percentage deviations from the trend by the formula

$$x_1 = 100.0\,(x - \exp\langle\ln x\rangle) / \exp\langle\ln x\rangle \qquad (1)$$

The averaging operator <...> represents the 60-point ramp filtering of the data.

Linear filtering is the same as a weighted average. For a time series variable $f(t)$ it is given by

$$\langle f(t_i)\rangle = w_1 f(t_i) + w_2 f(t_{i-1}) + \ldots + w_n\, f(t_{i-n+1}) \qquad (2)$$

For the 60-point ramp filter, n = 60 and

$$\begin{aligned} w_1 &= 119/1830 = 0.0650273\ldots \\ w_i &= w_1 - (3i-3)/1830,\ i = 2, 3, \ldots, 60 \end{aligned} \qquad (3)$$

Note that the filter coefficients decrease by a constant amount and eventually become negative, hence the name "ramp filter." See Morrison and Morrison (1997) for the general formula and a sketch of the derivation.

Applying equations (1) and (2) to the indices x, y, and z yields the percent deviations of the indices from the trend, denoted as x_1, y_1, and z_1. These points form a sort of donut (toroidal) shape distribution in 3-space. When lines connecting subsequent points are drawn, the gradual, irregular progression of the business cycle becomes obvious. However, there is nothing like angular momentum or even energy in the dynamics of the business cycle, so it may oscillate in one small region for months or even more than a year.

The final step consists of projecting the three-dimensional business cycle model onto a plane or other surface to get a model with just two parameters, a radius and a phase angle. To retain the integrity of the coincident indicator, we chose to restrict our choice of surfaces to planes passing through the y-axis. This may be suboptimal, but our philosophy is to make improvements in incremental steps.

Matrix notation provides an easy way to express this penultimate step

$$\mathbf{r}_2 = (x_2, y_2, z_2)^T,\ \mathbf{r}_1 = (x_1, y_1, z_1)^T \qquad (4.1)$$

$$\mathbf{r}_2 = \mathbf{G}\mathbf{r}_1 \qquad (4.2)$$

$$\mathbf{G} = \begin{bmatrix} \cos\gamma & 0 & -\sin\gamma \\ 0 & 1 & 0 \\ \sin\gamma & 0 & \cos\gamma \end{bmatrix} \qquad (4.3)$$

Of course, **G** is a very basic rotation matrix. The angle γ is restricted to the range from 0 to 90 degrees and evaluated by minimizing the sum of the squares of z_2. As a practical matter this was done by trial and error rather than nonlinear regression. It was less time consuming to make a number of runs of the transformation equations than to code the regression equations. A loop to compute the rms of z_2 was added to the code and displayed on the screen. The value determined was

$$\gamma = 53.2^\circ \qquad (5)$$

Note that this will weight the lagging indicator slightly more than the leading indicator; the angle would have to be 45° for equal weights.

To create the final polar coordinates, the new leading-lagging indicator was weighted by

$$x_3 = x_2 \div (\,|\sin\gamma| + |\cos\gamma|\,) \qquad (6)$$

Of course, $y_3 = y_2$; $z_3 = z_2$ was not changed because its only role is to have its rms minimized. This weighting was done to obtain values of the radial coordinate comparable with those of the previous two-index model and to eliminate phase angle shifts due solely to the change of scale along the new x_2-axis.

Polar coordinates in the x_3-y_3 plane are then obtained from

$$\begin{aligned} \rho_3 &= (x_3^2 + y_3^2)^{1/2} \\ \theta_3 &= \tan^{-1}(y_3/x_3) \end{aligned} \qquad (7)$$

These comprise the phase plane model of the business cycle. This new, improved model just replaces x_1 and y_1 of the former model with $y_3 = y_1$ and the new variable

$$x_3 = 0.42795\, x_1 - 0.57205\, z_1 \qquad (8)$$

Some economists have preferred to use the index of lagging indicators, inverted, instead of the index of leading indicators. Equation (8) is a weighted mean of the index of leading and the index of lagging indicators, inverted, to an approximation of the first order. The variables are percent deviations from the trend (think "differentials"), not the indices themselves, so the minus sign is all that is needed to specify "inverted," whatever may be the formula used for it, as long as its derivative is negative.

The basic concept for the improved model is that a weighted average of the leading index and lagging index, inverted, is better than either one alone. The geometric concepts in the model make it dynamically plausible. The expectation is that the improvement will be noticeable, but not dramatic.

4. Mathematical Modeling of Complex, Nonlinear Systems

Computer modeling of complex, nonlinear dynamic systems has been attempted many times as the machines progressed from huge, costly mainframes to even bigger and more expensive supercomputers. The personal computer now allows the average scientist, engineer, economist or forecaster, or even a self-taught amateur, to try his or her hand at the game. The degree of success has been underwhelming.

The modeling strategy that worked so well for classical celestial mechanics, and a few other areas in the physical and earth sciences, will fail in many other cases. There are few first approximations as good as Kepler's laws. Adding the mutual attractions of the moon and major planets produced a theory that served all practical and theoretical needs until the space age became mature. Now tidal effects have to be included for the most advanced missions and data analyses, so celestial mechanics is beginning to look more and more like economics.

This business cycle model provides something akin to Kepler's laws. But there are no equivalents to conservation of energy, linear momentum, or angular momentum. There is a stochastic inertia that keeps the cycle from making big jumps. An equally stochastic angular momentum makes the phase angle, θ_3, go forward most of the time, stall occasionally, and rarely go backward. Any random errors in the observations are swamped by biases and strongly correlated "filtered noise" behavior in the dynamics.

There is still a lot of serial correlation in the values of z_3. The strength of this signal could be reduced by using a curved projection surface rather than a plane, but the results, say z_4, the length of normals to the surface, would still be far from random noise. Any improvements in the phase plane model would be marginal. Models of complex, nonlinear systems reach a point of diminishing returns quickly. The orbits of the major planets comprise a large (60 variable), nonlinear system, but it is rendered simple by the weakness of the mutual gravitational attractions of the planets.

Orbital elements have a simple, geometric interpretation. GDP and the coincident indicators also display very simple dynamics in the zero-order model: exponential growth. Searching for leading and lagging indicators expresses a belief in the existence of a host of nonlinear feedbacks. Finding such indicators confirms the presence of such feedbacks, but the data are not nearly precise enough to resolve them. The indices, created through decades of work by many economists, provide only ill-defined aggregates of these many feedbacks.

Classical modeling consists of determining interactions, first the major ones. These provide a good zero-order model, like Kepler's laws. The secondary interactions, such as the perturbations of planetary orbits, provide a precise, practical model. Selected minor interactions, such as tidal effects and the variable rate of rotation of the earth, are needed only for some specialized applications.

Modeling in the age of *Chaos, a New Science*, as science writer James Gleick (1987) called it, is well illustrated by the development of this business cycle model. The initial data sets are a jumble of irregular cycles, or worse. Theories, where they do exist, seem to have nothing to do with the data. A new approach is needed.

The first job of the analyst is to create aggregates of the data that display some simple geometric patterns, or, failing that, are, to the greatest degree possible, amenable to forecasting. Creation of the three indices accomplished that step for the exchange economy of the USA. The indices and the GDP data display the exponential growth and the fact that there are significant, though irregular, deviations from that basic dynamical behavior.

This business cycle model provides a look at an aggregation of the more important feedbacks at work within the economy. The two-dimensional version (ρ_3 and θ_3) provides a model for visual, intuitive evaluation. The three-dimensional version [$\mathbf{r}_1 = (x_1, y_1, z_1)^T$] is quite suitable for modeling as a noise-driven difference (or differential) equation

$$\mathbf{r}_1(t+\Delta t) = \mathbf{A}[\mathbf{r}_1(t) + \mathbf{n}(t)]$$
$$\mathbf{n}(t) = \text{"noise"} \qquad (9)$$

The eigenvalues of the matrix **A** would comprise just about everything known about the dynamics of the economy, except the average growth rate. So the complete model would have only four parameters, a σ (standard deviation) for each component of **n** adds up to seven, plus a few initial conditions. That is not a lot, but it may be the best that can be achieved.

5. Is the "Improved" Model Better?

Not every change is for the better. Just a few years ago saw the introduction of New Coke. We can be sure that the company formulated the new product very carefully and tested it on a wide variety of consumers. Huge amounts of money were spent on advertising. But when the product arrived in the international marketplace, the world's cola drinkers sought out cans and bottles of Coke Classic and left the new offering sitting on the shelf.

Business cycle models are rather arcane, specialized products compared to soft drinks. Most of the basic materials are data collected by government agencies, with the rest coming from various private sources. The indices used to create this business cycle model were originally provided by the U.S. Department of Commerce, but the effort has been privatized and the work is being continued by The Conference Board in New York City. This model, like others, is a value-added product at the end of a long chain of supply.

The first question is whether this model is better than the one constructed from only two indices. The easiest way to approach that question is to look at the phase plane plots of both models for comparable time periods. A sample is given by Figures 1A, 1B, 2A, 2B, 3A, and 3B. The periods covered are roughly 1) 1976-1984, 2) 1983-1993, and 3) 1990-date. (They do overlap in time.) The "A" figure is the previous model and the "B" is the new model.

In large measure, the qualitative properties are retained. However, the phase angles are changed significantly in some cases, which was not true when the original model was recomputed from revised indices (Morrison and Morrison, 1998, 2000). The plots are more nearly circular in most cases. And sometimes they are rotated counterclockwise, especially in the most recent cycle. The identification of the official beginnings and ends of recessions is not improved very much.

Graphs of all the cycles will be available in a special Bulletin edition of the *Critical Factors* newsletter. Numerical tables of ρ_3, θ_3, and other parameters will also be available. These are omitted from this paper due to space limitations. More cycles could be determined from the earlier period truncated when The Conference Board did its first major revision of the indices. These results could be valuable in determining whether there have been changes over time in matrix **A** in equation (9) (and in its eigenvalues).

Our opinion is that the improved model is sufficiently better than the original one to justify the slightly increased effort needed to maintain it. The hypothesis tested in the original model would produce a uniformly circular "idealized" business cycle if the leading and coincident indices were both "perfect" (Morrison and Morrison, 1997). Incorporating the information in the lagging index makes most of the plots more nearly circular. We think that the significant counterclockwise rotation of the most recent cycle is an indication that this current cycle has been anomalous.

We had been continuing to make forecasts of the index of lagging indicators, even though we did not use them for anything. A better approach might be to forecast the noise vector $\mathbf{n}(t)$ and then use (9) to forecast the indices and hence the business cycle parameters, ρ_3 and θ_3. (It is easy to generate forecasts of the complete index from forecasts of percent deviations from the trend, if that is required.)

To date we have not compared either of these business cycle models or forecasts of them with the large, complicated econometric forecasting models that a number of sources produce. We do know that this model was much cheaper to develop and is much easier to maintain. One of our goals is to provide the best possible business cycle model and forecasts at a price that small businesses and individual investors can afford. (The CIA, DoD, the U.S. Treasury and the Federal Reserve can afford anything they think they need, but other agencies have limited budgets for reference materials and research.) A second goal is to provide a test of the modeling methodology described in Morrison (1991, Ch. 18, 19, 20) and summarized in Section 4 above.

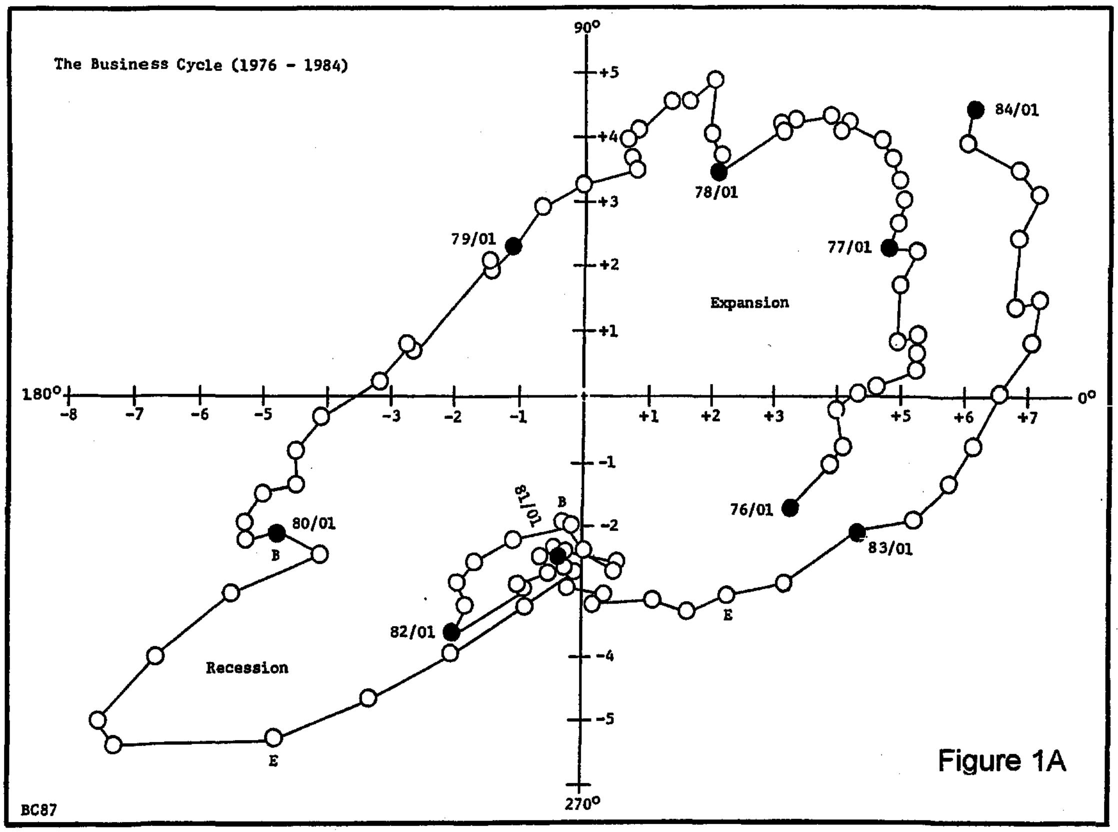


Figure 1A

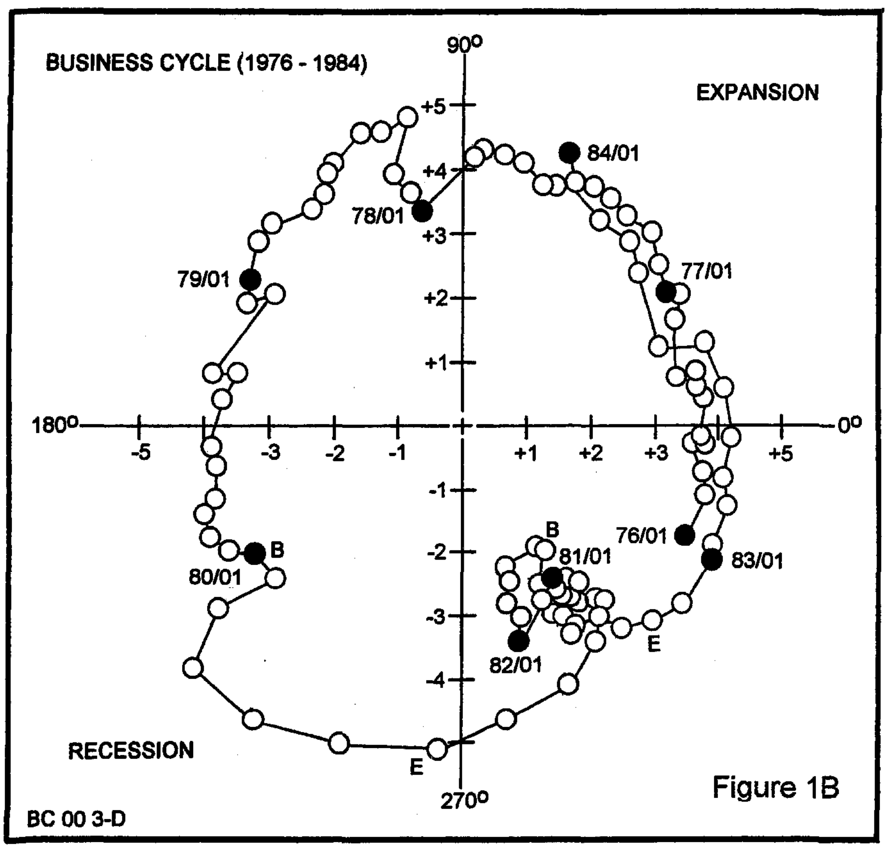


Figure 1B

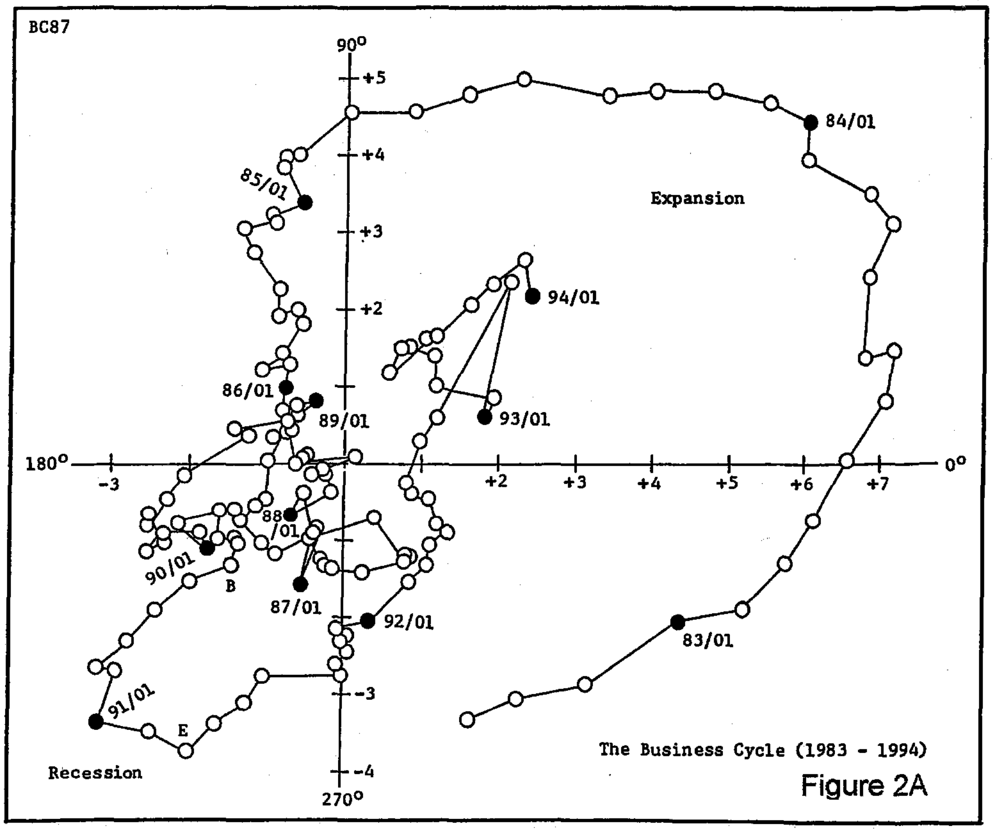


Figure 2A

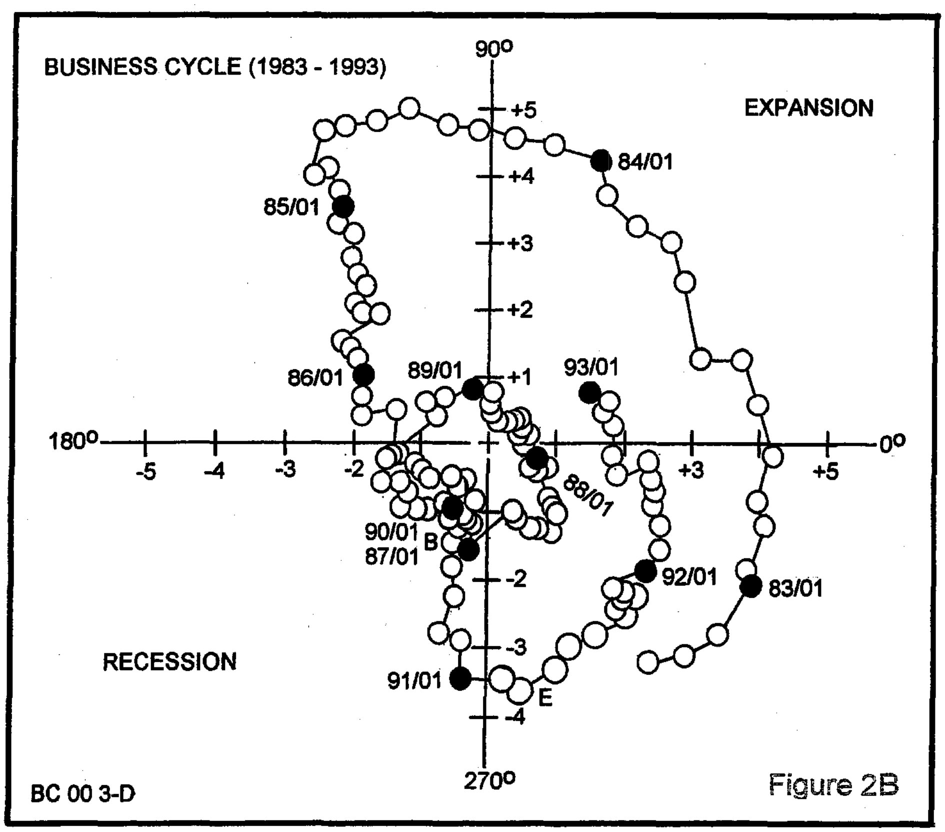
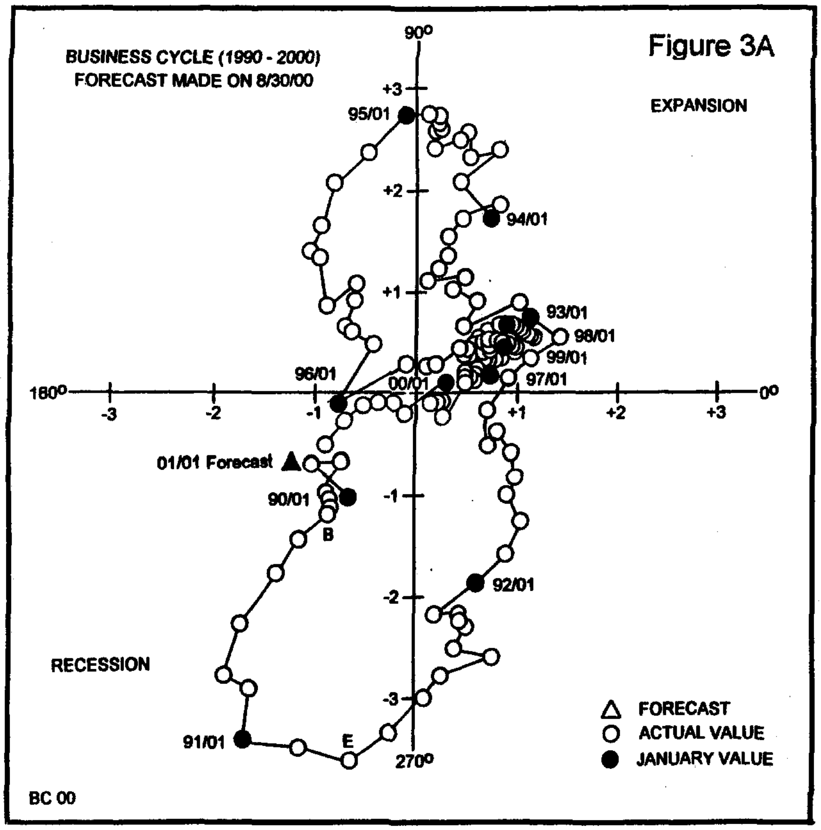
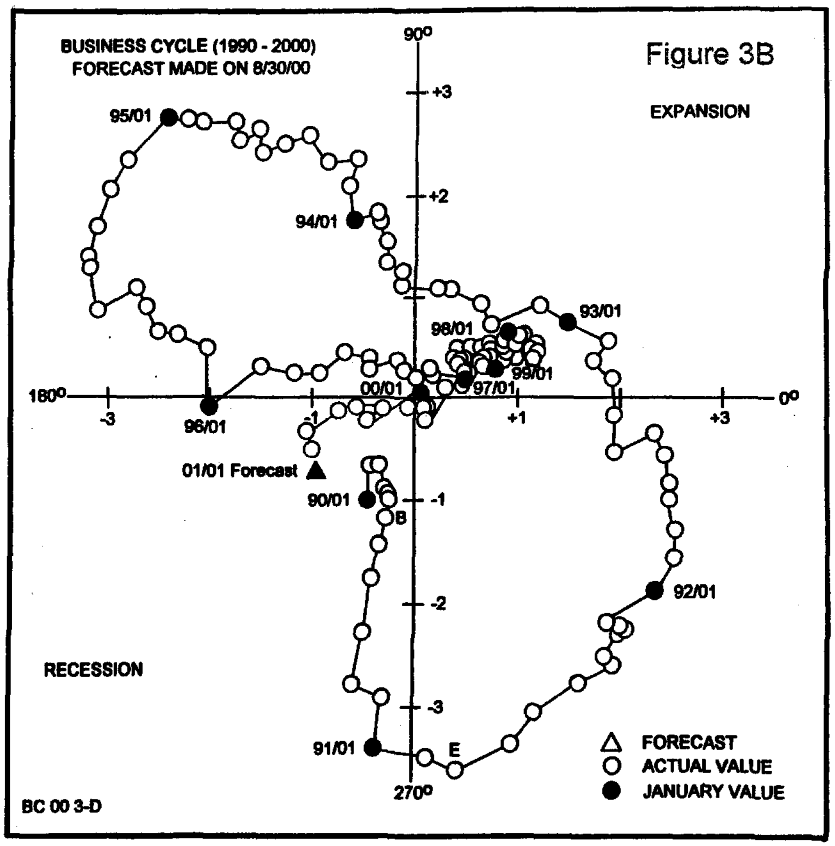


Figure 2B





Caption Figures 1-3.

For the "A" figures (previous model) the business cycle model is a phase plane plot of detrended leading and coincident indicators, as *x*- and *y*-coordinates respectively. Normal cycles follow a counterclockwise roughly ELLIPTICAL path with occasional stalls and reversals.

For the "B" figures (new model) the business cycle model is a phase plane plot of a weighted mean of the detrended index of leading and the detrended index of lagging indicators as *x*-coordinate and detrended coincident indicator as *y*-coordinate. Normal cycles follow a counterclockwise roughly CIRCULAR path with occasional stalls and reversals.

For "A" and "B" (both models) time is indicated along the cycle path. Expansions occur in the first quadrant (between *0°* and *90°*) and contractions in the third quadrant (between 180° and 270°). Other angles (second and fourth quadrants) denote transition periods. An "official" (National Bureau of Economic Research) beginning of a recession is indicated by a label "B" and an end by "E". Note that the 1976-1984 cycle had an official "double dip" recession.

The current cycle (1990-2000) includes a forecast. Note that the indicators used to construct the model are released about two months after the fact, so a forecast is needed to provide an estimate of the current value.

6. References

Gleick, J., *Chaos: Making a New Science*, Viking Penguin, New York, 1987.

Handbook of Cyclical Indicators: a Supplement to the Business Conditions Digest, 1984, US Dept. of Commerce, Bureau of Economic Analysis, Washington, DC.

Jordan, S.K., Self-consistent statistical models for the gravity anomaly, vertical deflections, and undulations of the geoid, *J. Geophys. Res.*, 77, No. 20, pp. 3660-3670, 1972.

Morrison, F., *The Art of Modeling Dynamic Systems: Forecasting for Chaos, Randomness, and Determinism*, Wiley-Interscience, New York, 1991.

Morrison, F. & N.L. Morrison, A phase plane model of the business cycle, *The 8th Federal Forecasters Conf. - 1996 & The 7th Federal Forecasters Conf. - 1994: Combined Papers & Proceedings*, US Dept. of Education, NCES 97-341, pp. 93-112, Debra E. Gerald, editor, Washington, DC, 1997.

Morrison, F. & N.L. Morrison, A revised phase plane model of the business cycle, *The 9th Federal Forecasters Conf. - 1997*, US Dept. of Education, NCES 98-134, pp. 115-122, Debra E. Gerald, editor, Washington, DC, 1998.

Morrison, N.L. & F. Morrison, An update on the business cycle, *The 10th Federal Forecasters Conf. - 1999*, US Dept. of Education, NCES 2000-017, pp. 45-52, Debra E. Gerald, editor, Washington, DC, 2000.

FORECASTING MODELS

Chair: Karen S. Hamrick
Economic Research Service, U.S. Department of Agriculture

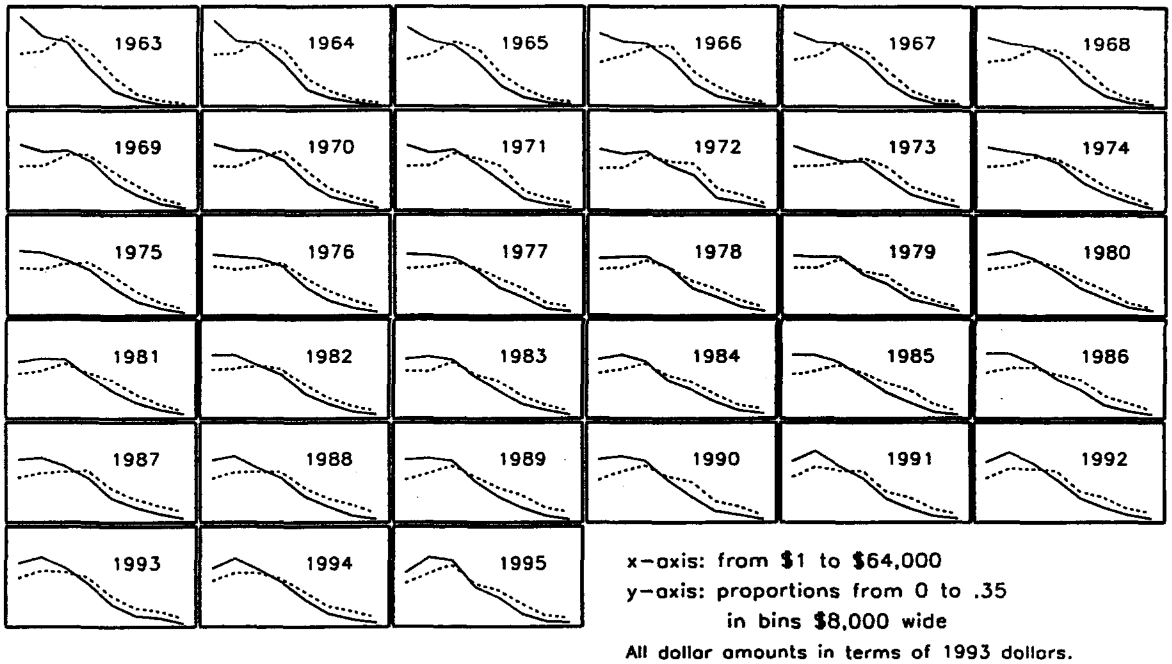
Forecasting the Convergence of the Rural and Urban Wage and Salary Earnings Distributions,
John Angle, Economic Research Service, U.S. Department of Agriculture

The Veteran Population Forecasting Model,
Allen Berkowitz and Stephen Meskin, U.S. Department of Veteran Affairs

Dynamic Programming of Forecasting Apparatus,
Elliot Levy, International Trade Administration, U.S. Department of Commerce

FORECASTING THE CONVERGENCE OF THE RURAL AND URBAN WAGE AND SALARY EARNINGS DISTRIBUTIONS

John Angle, Economic Research Service, USDA



Relative Frequency Distributions of Annual Wage and Salary Earnings
Rural Workers (solid curve) and Urban Workers (dotted curve)

Workers Aged 25 to 65

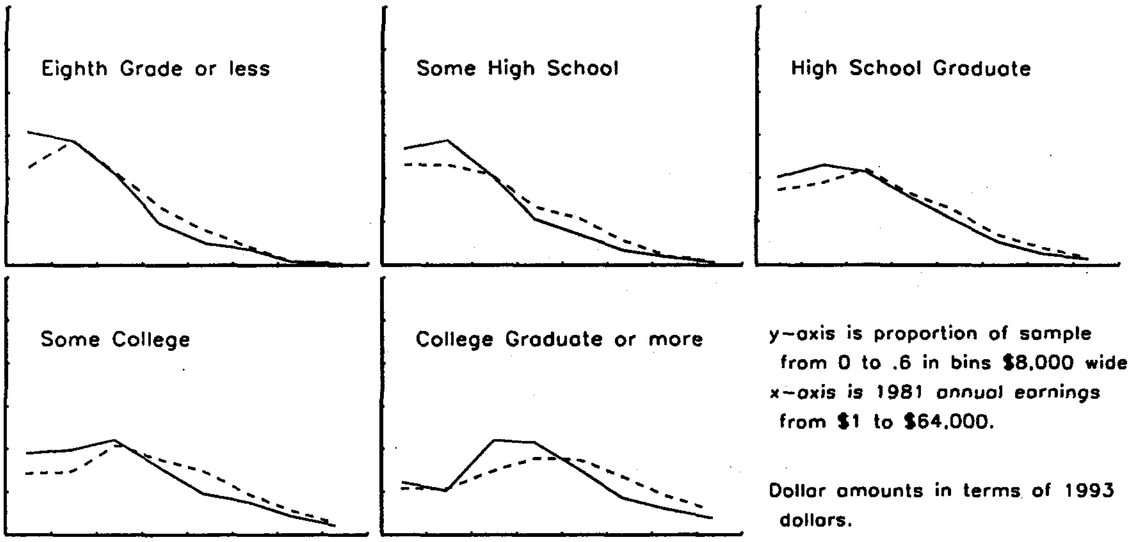
Source: March Current Population Survey

Figure 1

Introduction

Analysis of the gap between the rural[1] and urban distributions of annual earnings is important because the mean and median of the rural distribution have historically been well below the urban mean and median, while the rural proportion with low annual earnings has historically been higher than the urban. The Bureau of Agricultural Economics, the predecessor agency of the Economic Research Service within USDA, began studying rural economic well-being in the 1920's. Many rural residents judge the economic well-being of their communities in terms of the standard of urban economic well-being. It is this perception that drove net migration from rural to urban areas throughout the 19th and early 20th centuries but the rural/urban gap in the distribution of wage and salary earnings has shrunk in the late 20th century raising the possibility at least of the eventual *convergence of the two distributions* and the disappearance of the gap between rural and urban economic well-being. The foundation of a rural community's economic well-being is the distribution of wage and salary earnings of its residents. The urban

[1]This paper defines 'rural areas' as the set of nonmetropolitan counties. A nonmetropolitan county is a county not in a Metropolitan Statistical Area (MSA) as defined by the Office of Management and Budget (OMB). MSA's include core counties containing a city of 50,000 or more people or having an urbanized area of 50,000 or more and total area population of at least 100,000. Additional contiguous counties are included in the MSA if they are economically integrated with the core county or counties. The metropolitan status of every county in the U.S. is re-evaluated following the Decennial Census, with reclassification usually occurring at mid-decade. There has been a net decline in counties classified as nonmetro over the decades. However, the definition of nonmetro has remained more or less constant over the decades of data examined in this paper.



Relative Frequency Distributions of Annual Wage and Salary Earnings in 1981 by Level of Education

Rural: solid curve
Urban: dotted curve

Workers Aged 25 to 65

Source: March Current Population Survey

Figure 2

distribution is a relevant standard to evaluate the rural distribution of wage and salary earnings. Comparing the two distributions is a more comprehensive way of evaluating rural economic well-being than just examining particular descriptive statistics of earnings, e.g., the mean, the median, the proportion of low earnings, etc. A landmark of the literature in rural sociology on the gap between rural and urban economic well being, McGranahan's (1980) "The spatial structure of income distribution in rural regions", reviews a large literature which describes the rural/urban income gap primarily in terms of just two statistics, the median and the Gini concentration ratio, a measure of inequality. Knowledge of a distribution implies all the statistics of the distribution. The converse is not true so there is more information in the distribution than in any set of statistics that describe it.

Figure 1 graphs both the rural and urban distributions of annual earnings in each year from 1963 through 1995 inclusive. You can see in Figure 1 that in the mid-1960's the proportion of low annual earnings was much greater in rural areas than urban. Over the last thirty years, the distribution of annual wage and salary earnings[2] in rural areas became more similar to the urban distribution. Figure 1 shows the convergence of the rural and urban distributions of wage and salary earnings from the mid-1960's to the mid-1990's. Much of the difference between the two distributions in the 1960's was in the left tail, the proportion of workers with small wage and salary earnings. The left tail of the rural distribution was much higher than that of the urban distribution in 1963. It is evident from looking at the graphs of the two distributions over time from 1963 through 1995

[2]*The distributions in Figure 1 are estimated from the* 1964-1996 March Current Population Surveys (CPS). The Current Population Survey is a household survey with a large sample drawn and conducted by the U.S. Bureau of the Census. The smallest sample drawn in these years was more than 40,000 households. In March, CPS interviewers collect data on annual wage and salary earnings in the previous calendar year. The subset of the population that appears in Figure 1 is anyone, age 25 to 65, with at least $1 of wage and salary earnings in the previous calendar year. The minimum age of 25 is imposed on the sample to give students a chance to complete post-secondary education. The maximum age of 65 is imposed because many workers transition to retirement after that age. All dollar amounts in the data have been converted to 1993 dollars using the price deflators (CPI-U) of Table B-60 in Council of Economic Advisers (1998).

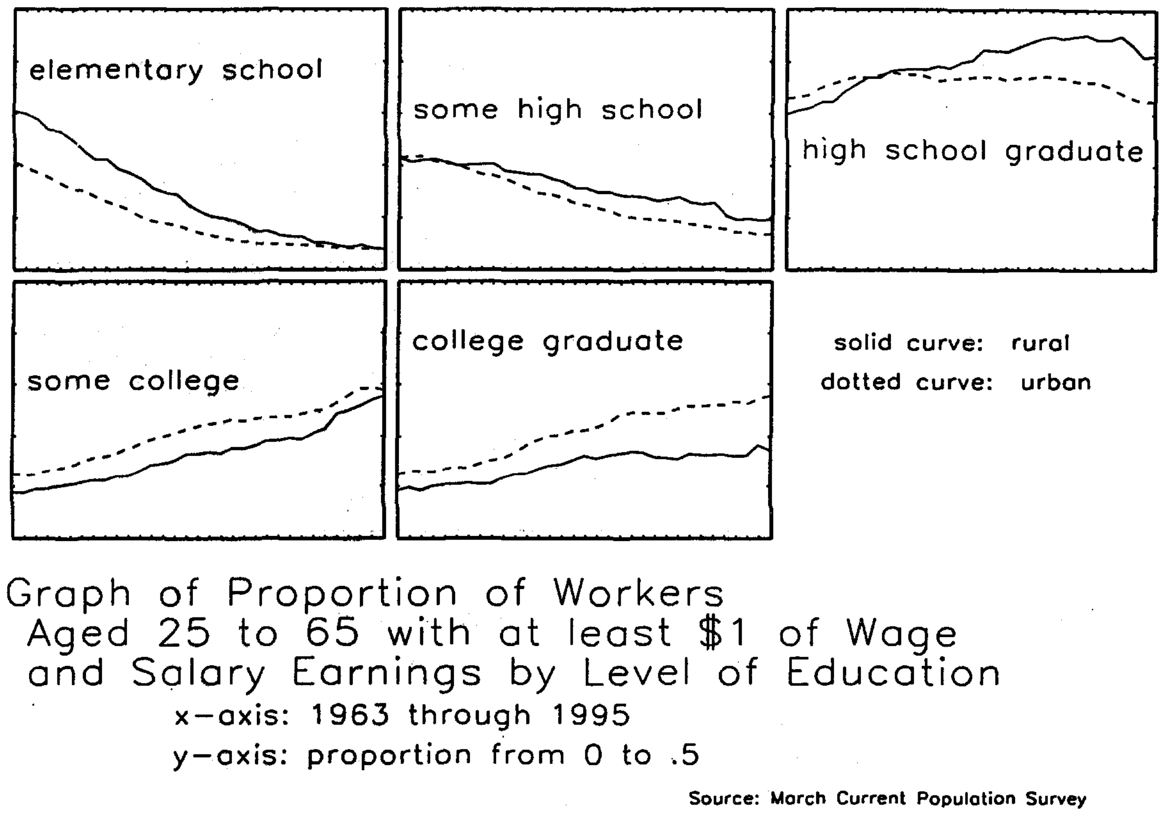


Figure 3

that the two left tails have grown closer together and that much of this convergence is due to the left tail of the rural distribution descending until it almost touches the left tail of the urban distribution. The rural distribution appears to have converged to the urban. This change represents progress because there is a smaller proportion of low earnings workers in urban than in rural areas.

Figure 2 shows the 1981 rural and urban distributions partitioned into five partial distributions of wage and salary earnings conditioned on education. Any other particular year in the data set would yield a similar result. The five partial distributions of Figure 2 when weighted by the proportion of workers at each education level add to the corresponding 1981 distribution in Figure 1. Notice that the lower the level of education in Figure 2, the higher the left tail of the distribution, i.e., the bigger the proportion of low earnings workers. Figure 2 shows that the shapes of earnings distributions conditioned on education appear to be similar in 1981 in rural and urban areas. Notice that the distributions in Figure 1 are shaped more like the distributions of workers with at least a high school diploma in Figure 2 than the distributions of workers without a high school diploma.

A Conjecture

Could it be that much of the convergence between the rural and urban distributions in Figure 1 is due to a decline in the proportion of workers without a high school diploma in both rural and urban areas but a greater decline in rural areas, erasing the distinctively higher left tail of the rural distributions in Figure 1 by the 1990's? If so, one might conjecture complete convergence to statistical indistinguishability of the two distributions.

One of the premises of this conjecture is supported by Figure 3, which shows declines in the proportions of workers without a high school diploma in rural and urban areas, with a particularly steep decline in the proportion of these workers in rural areas. By the mid-1990's, the rural proportion of workers with at most an elementary school education had plunged and almost converged to the low urban proportion. The urban proportion of the least educated

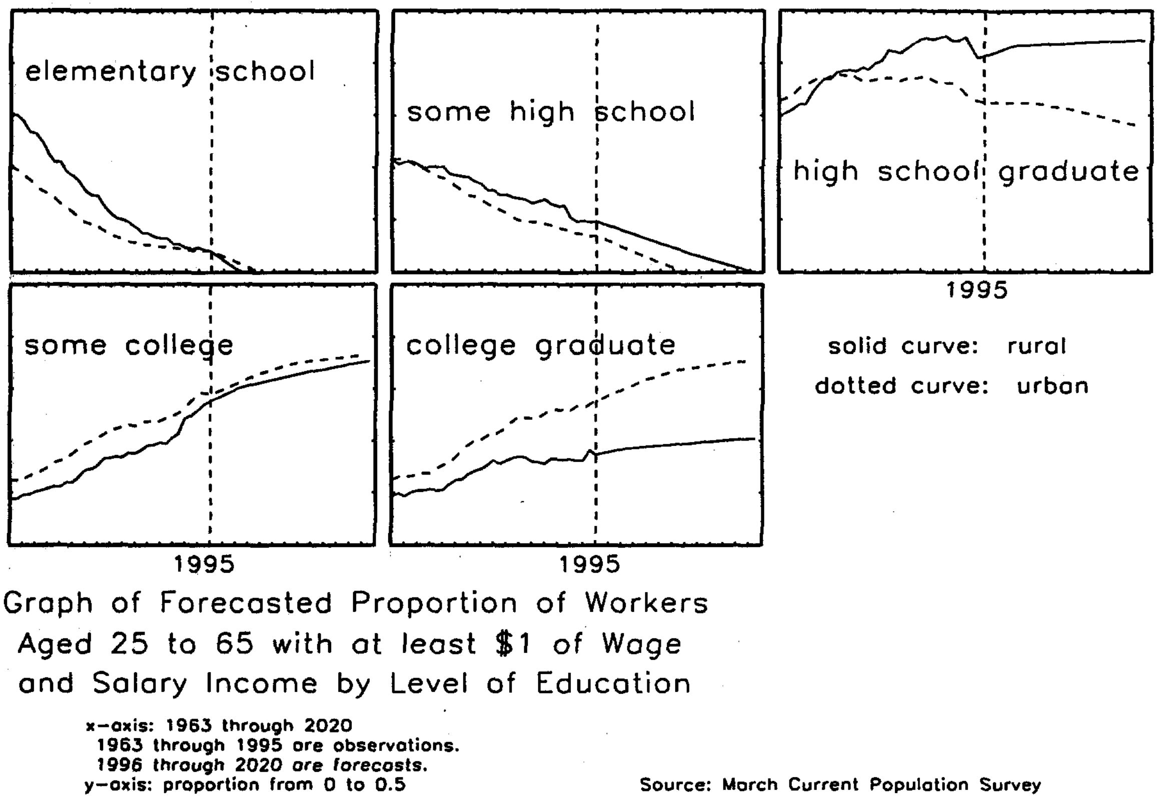


Figure 4

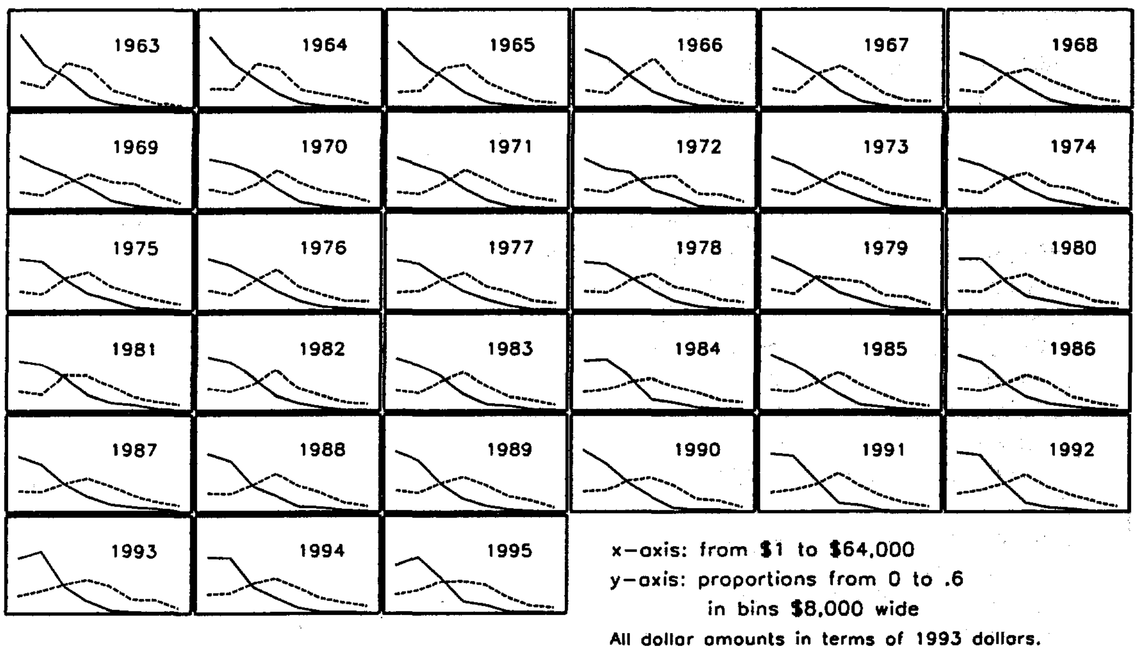
had fallen too, but not as far or as fast from its low 1963 value.

The economic history of the U.S. since the 18th century has been one of increasing integration and elimination of regional differences and barriers to competition. One might conjecture that this process will soon level differences between the rural and urban distributions of annual earnings. This paper attempts to estimate the time to convergence of the rural and urban distributions of annual earnings by forecasting the proportions of people at five major levels of education, the education levels in Figures 2 and 3. These proportions can be readily forecasted and are in Figure 4. The curves to the left of the dotted vertical line in Figure 4 are identical to the curves in Figure 3. These are the observed proportions of workers at each education level from 1963 through 1995. The curves to the right of the vertical dotted line in Figure 4 are forecasts. The method of the forecast of proportions at the two higher and two lower levels of education is to fit a straight line to a time-series of proportions. The forecast is the extrapolation of this line forward twenty-five years from 1996 through 2020 using the last observation, 1995, as the intercept. The middle education group, high school graduates, is forecast as 1.0 minus the sum of the other forecasted proportions. Then all the forecasted proportions are adjusted up or down to sum to 1.0 in each year. This adjustment introduces a non-linearity into the forecasts. The r^2 of each of the eight OLS regressions is given in Table 1.

The other condition of the conjecture is that, at least roughly, the shape of the conditional distribution, annual earnings conditioned on level of education, has not changed much from 1963 through 1995. This assumption can be examined on a rough basis by inspecting Figure 5 and Figure 6. Figure 5 gives the distribution of annual earnings of rural workers with at most an elementary school education and rural workers who are college graduates from 1963 through 1995. You can see in Figure 5 that while the shapes of the distributions of the least and most educated groups among rural workers changed somewhat, the basic shapes and the basic difference in shape between the distributions of the least and most educated persisted from 1963 through 1995. The same can be said about the comparable urban distributions in the same time period. See Figure 6.

Table 1. OLS regression results in forecasting proportions at four education levels, rural and urban, 1963 to 1995

rural/ urban	educat-ion level	regression coefficient	s.e.	r^2
rural	at most element-ary school	-0.008740	.000365	.95
rural	some high school	-.003950	.000117	.97
rural	some college	.005611	.000195	.96
rural	college graduate	.002735	.000195	.86
urban	at most elemen-tary school	-.005221	.000312	.90
urban	some high school	-.005162	.000156	.97
urban	some college	.005331	.000145	.98



Relative Frequency Distributions of Annual Wage and Salary Earnings for Rural Workers with Elementary School Educations (solid curve) and Rural Workers who are College Graduates (dotted curve)

Workers Aged 25 to 65 with Rural Residence

Source: March Current Population Survey

urban	college graduate	.005276	.000185	.96

So, intuitively, the idea of explaining the apparent convergence between the rural and urban distributions of annual wage and salary earnings by extrapolating a trend toward higher education levels and away from lower education levels especially in rural areas makes sense. The earnings distribution of the least educated workers is quite different from that of more educated workers. Rural areas had a much larger proportion of workers with at most an elementary school education in 1963 than urban areas. By 1995 workers with at most an elementary school education were almost gone from the rural as well as the urban labor forces. It is reasonable to conjecture that this trend produced the convergence between the rural and urban distributions and to conjecture that the continuance of this trend for workers with some high school education but no high school diploma will lead to near identical rural and urban earnings distributions. A precise measure of the dissimilarity of distributions, the "distance" between them, is needed to make a forecast of when the rural and urban distributions will converge.

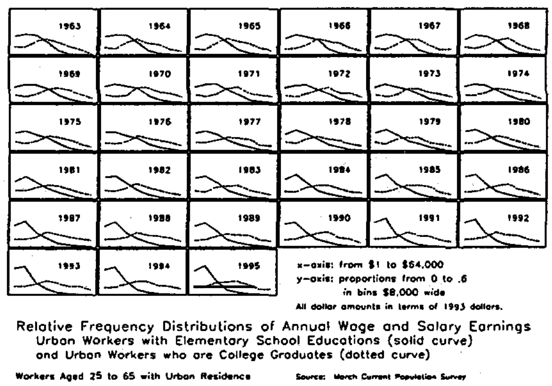


Figure 6

An Exact Measure of Distance between Distributions

Figure 1 does not provide a metric for the distance between the rural and urban unconditional distributions of wage and salary earnings. One of the best descriptors of the difference between two discrete distributions defined on the same set of relative frequency bins, as in Figure 1, is the symmetric entropy distance (Kullback, 1959:190), also called the symmetric Kullback entropy distance, the symmetric Kullback-Leibler distance, or the symmetric cross-entropy. The properties of this measure are discussed in Chapter 4 of Kapur and Kesavan (1992). The measure is defined between two distributions. The symmetric entropy distance between the rural and urban distributions, is, taking the relative frequencies of the rural distribution as p_i, and the relative frequencies of the urban distribution as q_i:

$$\sum_{i=1}^{I} (p_i - q_i)(\ln(p_i) - \ln(q_i))$$
$$= \sum_{i=1}^{I} p_i(\ln(p_i) - \ln(q_i)) + \sum_{i=1}^{I} q_i(\ln(q_i) - \ln(p_i))$$

The symmetric entropy distance is the sum of the asymmetric entropy distances between the two distributions. A symmetric entropy distance of 0.0 means that the two distributions are statistically indistinguishable, having the same relative frequencies in each bin. Figure 7 shows that from 1963 through 1979, the symmetric entropy distance between the rural and urban annual distributions of wage and salary earnings plunged to about .05. You can see in Figure 1 that a symmetric entropy distance of about .05 (in 1979) means that the two distributions partially overlap in their right tails and central masses although they are clearly distinct in their left tails. The standard errors of the relative frequencies are quite small given the enormous sample sizes and considering them is not useful in interpreting Figure 7.

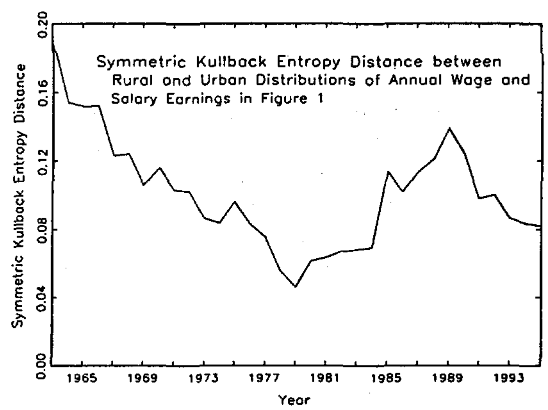
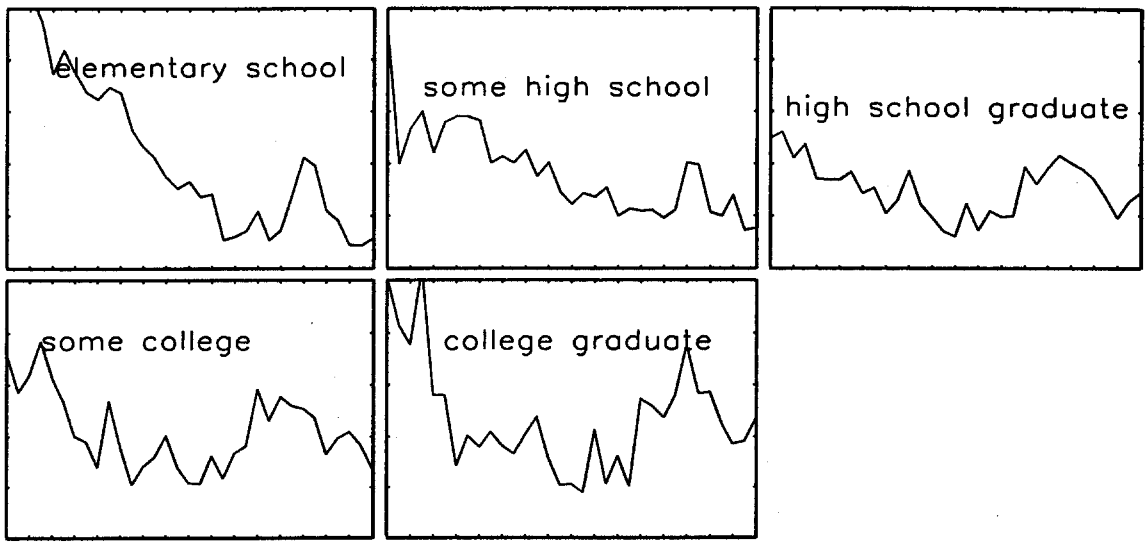


Figure 7

Figure 7 tells a story that is only partially apparent in Figure 1. True there is the dramatic convergence between the rural and urban distributions from 1963 through about 1979 that is apparent in the inspection of Figure 1. But there is also a reversal of this trend during the 1980's that is much more difficult to discern in Figure 1. This divergence however is transient and by 1995 the two distributions are back to a symmetric entropy distance of about .08. Maximum divergence during this transient episode occurred in 1989. You can see in Figure 1 that the state of



Symmetric Kullback Entropy Distance between
Rural and Urban Conditional Distributions, Annual
Wage and Salary Earnings Conditioned on Education
x–axis: 1963 through 1995
y–axis: symmetric entropy distance from 0 to .2 (same scale as in Figure 7)
Source: March Current Population Survey

Figure 8

divergence in 1995 is not great. The two distributions were close in 1995. However, the episodes of convergence, divergence, and re-convergence in the 33 years from 1963 through 1995 do not lend themselves to a forecast of whether the two distributions will become indistinguishable in terms of the symmetric entropy distance.

The symmetric entropy distance in Figure 7, while not a simple weighted sum of the symmetric entropy distances between the rural and urban partial distributions of the conditional distributions, wage and salary earnings conditioned on education, can be greatly affected by these distances, particularly if the partial distributions are at least somewhat similar to each other, as Figures 2, 5, and 6 suggest. The weights referred to here are the proportions at each education level in rural and urban areas. If these remain constant over time, one would expect divergence between the rural and urban partial distributions at a particular level of education to increase the distance between the unconditional rural and urban distributions in Figures 7 and 1. As Figure 3 shows, the weights are changing. The proportion of workers, rural and urban, with at most an elementary school education is declining. The decline is at different rates though, faster for the rural population than the urban population. Conversely, the proportions of workers at higher levels of education are increasing. The rural and urban proportions at the 'some college' level are increasing apace. At the highest level, 'at least college graduate', both proportions are increasing but the urban proportion is increasing faster than the rural proportion.

Figure 8 shows that the symmetric entropy distances between the rural and urban partial distributions of groups at high levels of education have little trend toward convergence after 1981. In fact these show the most divergence after 1981. Figure 9 shows that the higher the level of education, the greater the divergence after 1981. The average rank of distance between the partial distributions in Figure 8 over the five levels of education (with a rank of 1 meaning the closest and 5 the most distant) from 1982 to 1995 is:

1.429 at most elementary school
2.286 some high school

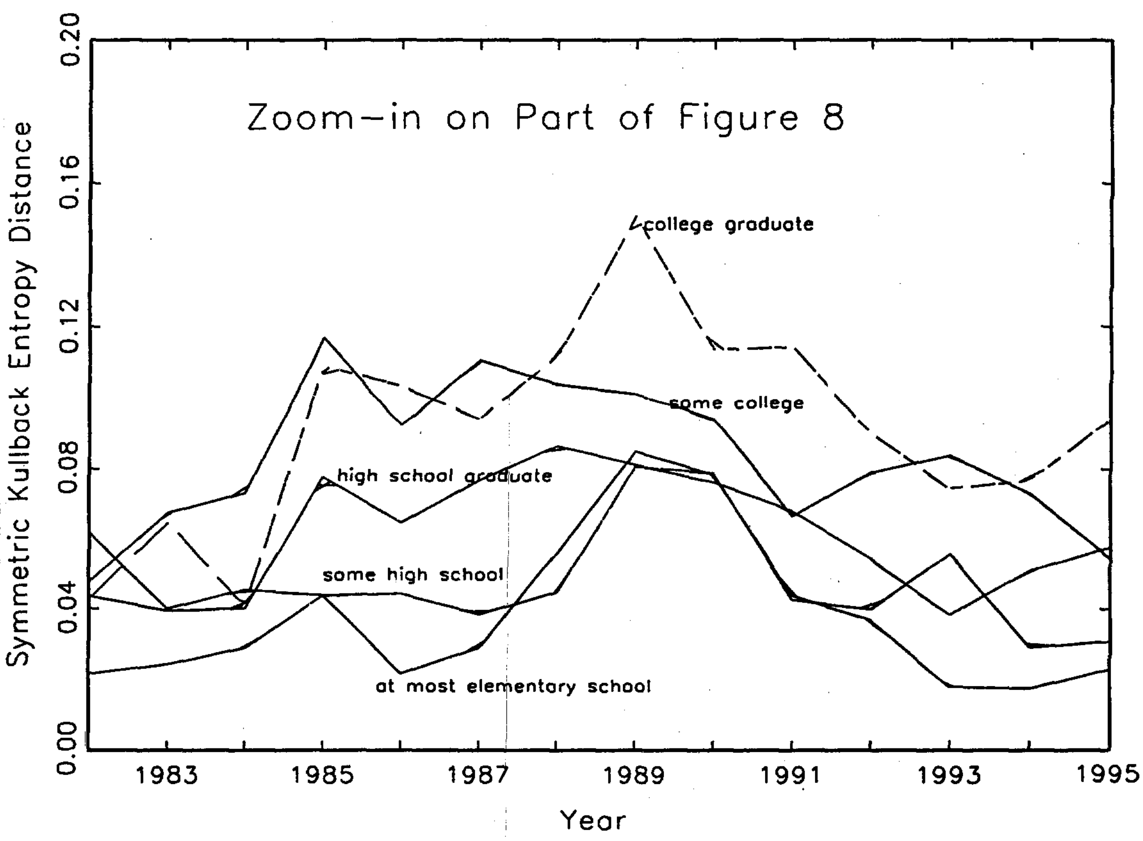


Figure 9

2.714 high school graduate
4.214 some college
4.357 college graduate.

See Figure 9, which zooms in on Figure 8 from 1982 to 1995. Figure 9 shows that the divergence between the rural and urban distributions in the 1980's was related to education level: the higher the level of education the greater the divergence.

Figure 9 should be compared to Figure 10, the comparable graph for the years 1963-1975. These are the years of the most rapid convergence. There is no clear ordering by level of education. The rankings of education groups in terms of the symmetric entropy distance between the rural and urban partial distributions of Figure 8 are:

4.923 at most elementary school
3.077 some high school
1.385 high school graduate
2.385 some college
3.231 at least college graduate.

Figure 10 shows that during the main episode of convergence the inverse ordering of distances between rural and urban partial distributions by level of education did not obtain.

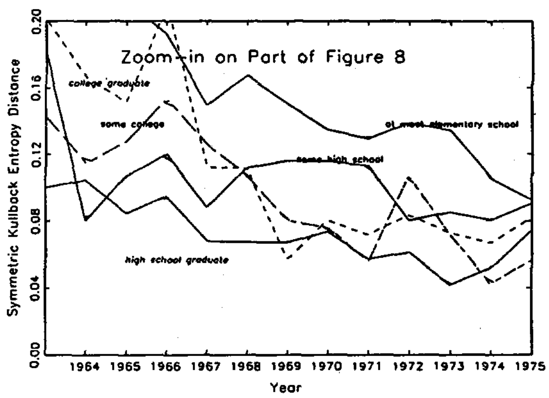
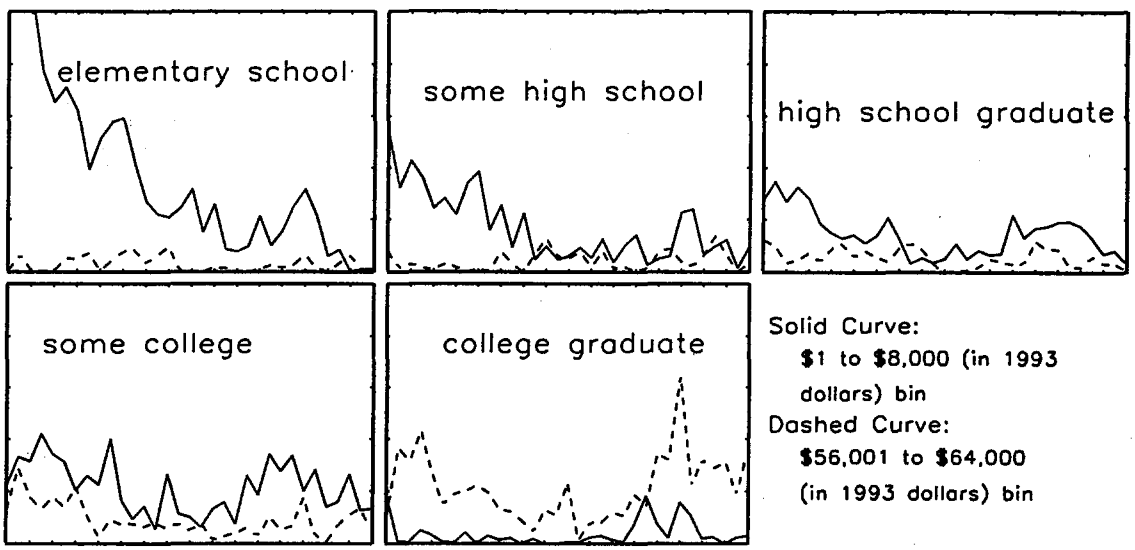


Figure 10

The Forecast

A visual inspection of Figure 1, the distributions of annual wage and salary earnings in terms of 1993 dollars from 1963 through 1995, in rural and urban areas shows that by 1995 these distributions had substantially converged. This paper raises the question of whether this convergence can be extrapolated into the future to the point that one can say that the distributions are statistically indistinguishable. Much of the convergence between the two distributions is due to the higher left tail of the rural distribution coming down to overlap that of the urban distribution. This movement is not just geometry. The left tail is the proportion of people in the relative frequency bin of the smallest income range, from $1 to $8,000 in terms of 1993 dollars. It is good news that the proportion of rural workers earning more than that has increased substantially.

Figure 2 shows that the left tail of an earnings distribution has a strong relationship to a worker's level of education. The higher the level of education, the lower the left tail, i.e., the smaller the proportion with the smallest annual earnings. There has been substantial progress in rural education, i.e., rural areas catching up to urban areas in school completion rates The effectiveness of rural schools in the last four decades of the 20th century has improved as well. See McGranahan and Ghelfi (1991) and Gibbs, Swaim, and Teixera (1998). So it makes sense to hypothesize that a disproportionate decline in the least well educated in the rural labor force is what caused the convergence. Figure 3 shows that there have been steady declines to almost the same tiny proportion in both rural and urban areas of workers with at most an elementary school education. The decline has been steeper in rural areas. The proportion of the next higher education level distinguished in Figure 3, 'some high school' shows that the rural proportion has never been much higher than the urban proportion and that both have declined, although not as quickly as the proportions of workers with at most an elementary school education. The two highest levels of education distinguished, 'some college' and 'at least college graduate' have shown increases both rural and urban. At the 'some college' level the proportions have increased apace. It is in the highest category 'at least college graduate' that the urban proportion has increased more rapidly than the rural proportion and is opening a lead. The proportions at high and low levels of education change in a near linear way. They appear to be readily forecastable via linear extrapolation. See Figure 4 for the 25 year forecast from 1996 through 2020. The proportion 'at most elementary school' both rural and urban is almost zero already. The



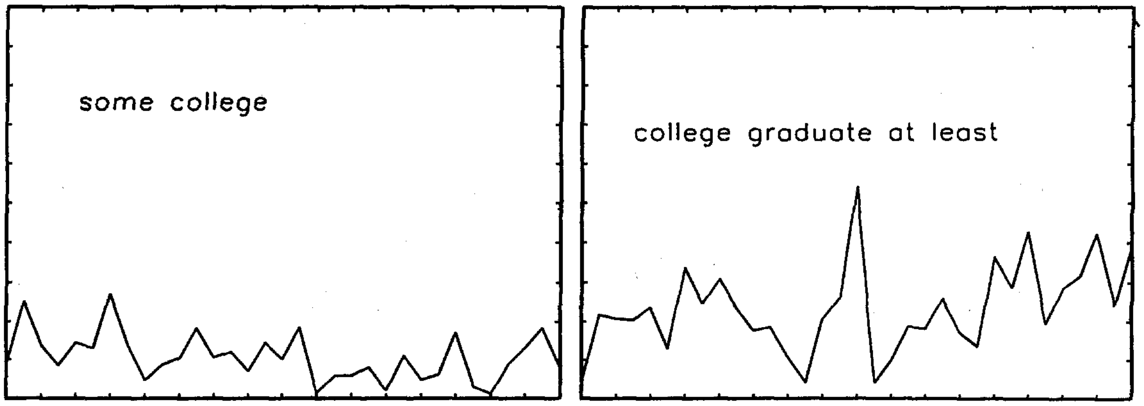
Symmetric Kullback Entropy Distance between Bins of Lowest and Highest Earnings in the Rural and Urban Conditional Distributions, Annual Wage and Salary Earnings Conditioned on Education

x-axis: 1963 through 1995

y-axis: symmetric entropy distance from 0 to .1 (2x scale of Figures 7 and 8)

Source: March Current Population Survey

Figure 11



Proportion of Symmetric Kullback
Entropy Distance due to contribution of bin
of largest incomes, i.e., $56,001 to $64,000

x-axis: 1963 through 1995
y-axis: proportion from 0.0 to 1.0

Source: March Current Population Survey

Figure 12

proportion 'some high school', urban and rural, is forecast to be almost zero by the year 2020. The data and this extrapolation procedure shows a widening gap by 2020 at the highest level of education between the rural and urban proportions, although both are increasing.

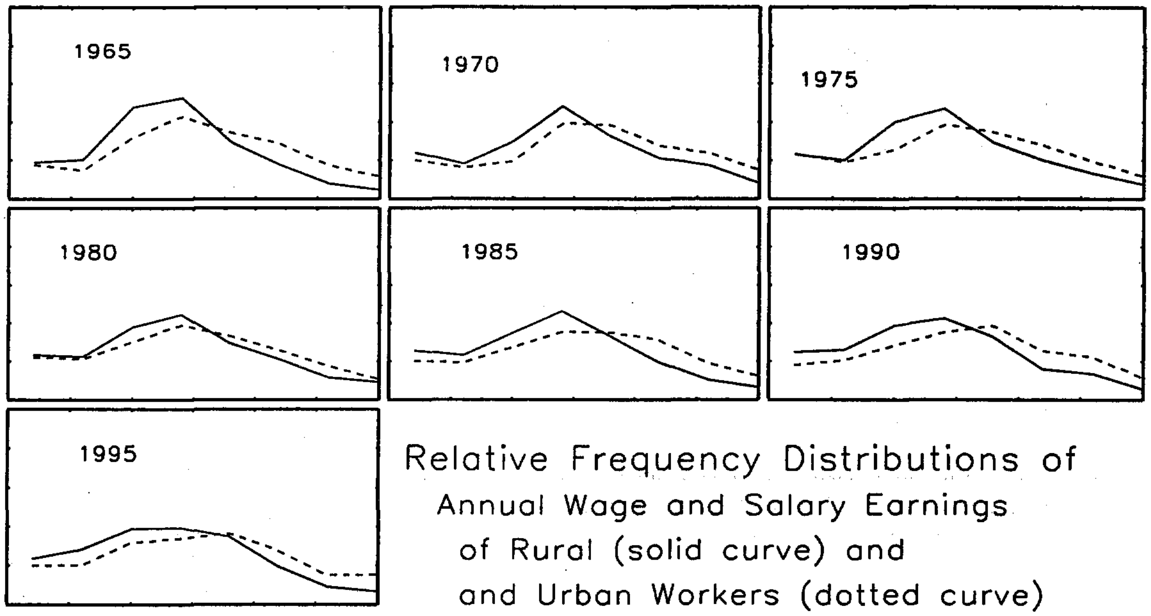
An exact measure of the difference between distributions is needed to understand and forecast convergence. The best measure is the symmetric entropy distance. The symmetric entropy distance between the rural and urban distributions of annual earnings is given in Figure 7. Indeed it shows substantial convergence between 1963 and 1995 but it shows something else not as readily discerned in Figure 1: a period of divergence following 1979, the year of maximum convergence. The year of maximum divergence was 1988, which was followed by reconvergence. So there is not a uniform convergence between 1963 and 1995, i.e., no uniform, incremental trend to simply extrapolate. Figure 7 shows that forecasting the future of rural and urban convergence in earnings distribution is inherently difficult.

A possible way around the difficulty with forecasting an aggregate is to decompose it to see if the components are more readily forecastable. Figure 8 does this for education levels. Figures 5 and 6 show that the relationship between education level and the shape of the earnings distribution is fairly stable over time. Figure 4 gives a credible forecast of the rural and urban proportions at each of the five levels of education distinguished. Figure 8 gives the distance between each partial distribution of the conditional distribution, annual earnings conditioned on education, rural and urban. Figure 8 shows several distinct patterns. First and most clearly, not only is the proportion of workers with only an elementary school education headed toward zero, so is any rural/urban difference in their distribution of earnings. There is also a clear pattern of convergence between the rural and urban partial distributions of the next two higher levels of education, 'some high school', and 'high school graduate'. The three lower levels of education show some divergence in the 1980's, but most of this divergence occurs at the 'some college' and most clearly at the 'college graduate' level. In fact, if you overlap the time-series of the rural urban distance between the partial distributions from 1980 on, as in Figure 9, you see that the higher the level of education, the greater the divergence between the partial distributions. The earlier period of convergence showed no comparably clear ordering in terms of education level.

Figure 11 shows that it is the symmetric entropy distance between the left tail of the least well educated groups that greatly decreased between 1963 and 1980, the period of convergence but that it is the symmetric entropy distance between the right tails, the largest earnings bin ($56,001 to $64,000 in 1993 dollars), of the most educated group that greatly increased during the divergence of the 1980's. Figure 12 shows that as a proportion of the entropy distance between the rural and urban partial distributions, Figure 8, the contribution of the right tail, the rightmost bin, is increasing among the most educated workers, that is, the educational group whose urban proportion is outstripping the rural proportion.

You can see in Figure 13 that the convergence between the rural and urban distributions of annual earnings of the most educated group, workers who have completed at least four years of college, has not been substantial. It looks as if in 1995 there is divergence in the right tail. This visual inspection of the relative frequency distributions confirms the inference drawn from Figures 11 and 12.

So it appears that the convergence between the rural and urban distributions of annual earnings around 1979 was the result of a) the convergence between the rural and urban proportions of workers with at most an elementary school education to b) almost zero and c) a convergence between the rural and urban partial distributions of the earnings of workers with at most an elementary school education.



Relative Frequency Distributions of Annual Wage and Salary Earnings of Rural (solid curve) and and Urban Workers (dotted curve) who are at least College Graduates

x-axis: from $1 to $64,000
y-axis: proportions from 0 to .6 in bins $8,000 wide
All dollar amounts in terms of 1993 dollars.

Workers Aged 25 to 65 Source: March Current Population Survey

Figure 13

There were other trends afoot in the period 1963 to 1995. There was the trend toward a greater proportion of workers in the two higher education level groups. The rural and urban proportions of 'some college' workers have both been increasing steadily. While both rural and urban proportions of workers who have graduated from college have been increasing, the urban proportion has been increasing at a faster rate than the rural proportion. Not only have the rural and urban proportion of workers who are at least college graduates been pulling apart, their partial distributions of annual wage and salary earnings have been diverging irregularly since 1979 as well. This pattern of divergence between the rural and urban distributions of the most educated is clearest in the extreme right tail of the distributions, the distribution of workers over large incomes. The urban proportion is larger than the rural proportion and the difference in the right tail is becoming larger slowly.

This paper intended to make a forecast based on a past trend. The trend is the decline in the rural and urban proportions of workers with only an elementary school education to almost zero from 1963 through 1995. The premise of the forecast is that this trend accounts for the convergence of the rural and urban distributions of wage and salary earnings. While this premise is substantially correct, it is not a basis to make a forecast from. After 1979 two other trends affected the distance between the rural and urban distributions of wage and salary earnings. One trend is a divergence in the rural and urban proportions of the most educated in the labor force. The urban proportion is accelerating away from the rural proportion. The other trend is a divergence between the rural and urban wage and salary distributions of the most educated group distinguished in this study, those with at least a college degree. These two trends may eventually affect enough people to cause a substantial divergence between the overall rural and urban distributions of annual wage and salary earnings. However, as of 1995 the divergence between the rural and urban distributions of the most educated was sufficiently weak and involved sufficiently few workers that it is premature to forecast in the year 2020 a divergence between the overall rural and urban distributions on this basis. However, it can be confidently forecast on the basis of these trends that the overall rural and urban distributions are unlikely to converge between 1996 and 2020 more closely than they were at their point of closest convergence in 1979.

Conclusions

In the past, the rural and urban distributions of annual earnings differed because the rural proportion of the least educated workers was substantially greater than the corresponding urban proportion. Also, the rural and urban distribution of the annual earnings of such workers differed in that there were proportionately more low earnings in the rural distribution. In the future it appears that the rural and urban distributions of annual earnings will differ largely because the urban proportion of the most educated is greater than the rural proportion and increasing more quickly. Also, the urban distribution of the annual earnings of those with at least a college degree differs from the rural distribution in that its right tail is thicker. The rural and urban distributions of the earnings of workers with at least a college degree do not appear to be converging. The simplest explanation of this divergence is that not only is the urban proportion of workers who are at least college graduates increasing faster than the rural proportion but that the urban proportion of the more educated within this highly educated group is increasing faster than the rural proportion.

References

Council of Economic Advisers. 1998. **Economic Report of the President.** Washington, DC: U.S. Government Printing Office.

Gibbs, Robert, Paul Swaim, and Ruy Teixera. 1998. **Rural Education and Training in the New Economy** Ames, Iowa: Iowa State University Press.

Kapur, Jagat and H.K. Kesavan. 1992. **Entropy Optimization Principles with Applications** Boston: Academic Press.

Kullback, Solomon. 1959. **Information Theory and Statistics**. Mineola, NY: Dover.

McGranahan, David. 1980. "The spatial structure of income distribution in rural regions". **American Sociological Review** 45:313-324.

McGranahan, David and Linda Ghelfi. 1991. "The education crisis and rural stagnation in the 1980's". Pp. 40-92 in Richard Long (ed.) **Education and Rural Economic Development.** Washington, DC: Economic Research Service.

The Veteran Population Forecasting Model

Allen Berkowitz and Stephen Meskin
Office of the Actuary
Office of Policy and Planning
U.S. Department of Veteran Affairs

Introduction

The purpose of this paper is to describe improvements in the methodology used to develop projections of the veteran population by the Department of Veterans Affairs.

- **Description of the Veteran Population**

  As of July 1, 1999 the estimated number of veterans living in the United States and Puerto Rico stood at 24.1 million. This includes 8.1 million Vietnam era veterans, and, on the other extreme, approximately 3,000 living World War I veterans. At the same time, the estimated median age of veterans was 58.4 years, with 38% of the total projected to be over the age of 65. Female veterans were estimated to number 1.2 million.[1] Statistical Appendix, FY 1999 Annual Accountability Report, Department of Veteran Affairs)

  As opposed to estimates, the number of veterans actually receiving compensation for service-connected disabilities as of July 1, 1999 was 2,668,186 and those receiving pension benefits due to low income and total disablement was 367,588. The number of veterans enrolled in the VA Health Care System is 4,175,833. It should be noted that this is not a discrepancy from the overall totals as only a small percentage of veterans are entitled to compensation and pension benefits. Furthermore, although all veterans are currently eligible for VA health care, only a small percentage of veterans utilize the veterans health care system. One possible reason is a veteran may have other health coverage through their current employment.

  The statistics concerning the size and characteristics of all veterans (as opposed to beneficiaries) must be estimated each year with the exception of the decennial year when they can be obtained directly from the U.S. Census.

  Furthermore, projections for a period of thirty years into the future are desired for the total population, and for specific classes of beneficiaries in order to support planning and budgeting of VA resources.

- **Users of Population Statistics**

  OMB, Congress, Veterans Service Organizations, DoD, DoL, and State Veterans Directors within state governments are all external users of VA statistics. Internal to VA, population data are used by the Veterans Health Administration, the Veterans Benefit Administration, the National Cemetery Administration, and several other planning and budgeting divisions. The Veterans Benefits Administration, for example, uses the projections of the number of separations from the military to estimate compensation and pension workloads and expected expenditures for education benefits. Projections of veterans by locality are used by VA's National Cemetery Administration to determine VA cemetery development priorities. The VA Health Administration also uses veteran population data at the local level for capital planning purposes in the location of new health facilities and in determining its market share and potential for expanding enrollment.

  The importance of a customer focus by agencies throughout the federal government has increased in recent years. Veteran population numbers are more important, particularly at a disaggregated level and particularly if they can be produced by a parameter driven model that can be "upgraded" easily when new information becomes available.

Background

- **Framework**

The current projection model (VETPOP) was last run in 1993; it provided estimates for 1993, the base year, and projections of the veteran population, separations, deaths, and interstate migration through 2020. The base year population was estimated by starting with the decennial census information on veterans, adjusting for assumed deaths and migration in the intervening years, and adding known separations over the same period of time.

To project separations from the military, years of service dependent separation rates were developed from historical trends.

- **Specific Methodology**

The specific projection method used a variation of the Cohort Survival Rate Method to project the veteran population. (Figure 1). Specifically:

$$P^{a,g,p}_{t+1} = P^{a,g,p}_{t} + B^{a,g,p}_{t,t+1} - D^{a,g,p}_{t,t+1}$$

where:

P_t = veteran population at time t,

$B_{t,t+1}$ = separations from active duty military in the period t to t+1

$D_{t,t+1}$ = veterans' deaths during the period t to t+1; and

a = single year of age,
g = gender,
p = period of service

To project the veteran population at the state level, baseline information was obtained for the distribution of the veteran population by age and gender. Projected separations and mortality was applied at the state level. Net interstate migration rates were used to project the movement of veterans from one state to another. The projections of net interstate migration rates are based on data for the civilian population obtained from the Census Bureau.

A more detailed explanation of this model and the methodology used in projecting separations can be found in Sorensen[2,3].

Model under Development

- **Framework**

The new veteran population projection model (VPM) will have two main objectives. First, as in the prior model it will estimate and project the number of veterans by age, gender and period of service at the national, and state levels. County-level demographics will be estimated outside of the model. Second, it will estimate and project the number of veterans, surviving spouses, and surviving dependents that are eligible for, apply for, and utilize the following VA programs:

- Pension
- Compensation
- Health Care
- Vocational Rehabilitation
- Home Loan Guarantees
- Burial
- Education Benefits
- Internment in National Cemeteries

The theoretical framework for producing the national estimates in the new veteran population projection model is similar to the original model[4]. The base population year must be established and the cohort survival rate equations applied. There are several key differences in the methodologies used in the two models with respect to:

1. The method used to establish the baseline population.
2. The disaggregation of the veteran population in the baseline and in the projections to specifically address

tracking beneficiary classes (compensation, pension, non- disabled veterans).

3. The determination of the appropriate mortality table to apply to the veteran population.
4. The methodology used to project separations from the military.
5. The use of interstate migration rates uniquely established for the veteran population.
6. Enhancements to the dissemination of model outputs.

Each of these differences is discussed in reference to the diagram shown in Figure 2.

- **Specific Methodologies**

The method used to establish the baseline population

The VETPOP model established the baseline by aging the veteran population as of April 1, 1990 up to the date of the baseline estimate by applying mortality rates, and adding known separations from the Defense Manpower Data Center (DMDC) data files for the period from 1990 to the baseline date. The new model uses census data (from 1990 Census) only for the pre-Vietnam population. It then utilizes the DMDC data, that, when matched with internal VA Compensation and Pension (C&P) data, provides additional information on disability status and type of benefit received by classes of beneficiaries, for the period from May, 1975 to September, 1999.

Tracking of beneficiary classes

The new model projects the veteran population by distinct sub-populations of disabled and non-disabled. This imposes an additional methodological requirement of developing transition probabilities between disability classes. Beneficiaries who are survivors of deceased veterans are also tracked by maintaining a deceased record through the projection period.

Selection of mortality tables

For living veterans, two sets of mortality tables were used. One table, for healthy veterans, was derived from the mortality experience reported by the DoD Office of the Actuary. The second table was developed for the disabled population, based on actual experience with the VA Compensation and Pension programs. This table is further refined to distinguish between veterans with less than 40% combined disability ratings and higher levels. The number of separated veterans is adjusted for mortality from the time of separation up to the base year.

In developing an estimate of the number of deceased veterans prior to 1990 (to track survivors and dependents), mortality rates developed by the Office of the Actuary of the Social Security Administration were used. These rates are available in the Actuarial Study No. 107, Life Tables for the United States Social Security Area 1900-2080.

Projection of separations

Projections of the number of separations by year and age are provided by the Defense Department's Office of the Actuary. The model requires a further disaggregation of separation projections by state and gender. Four years of historical data from the period 1995 to 1998 were used to establish the state distributions of separations. Initially, the percentages by state are assumed to be constant over the projection period. In the past ten years, the number of female enlistments, in the military has increased significantly. This will be reviewed in future projections. An analysis of this impact on the percentage of female separations was used to project the change in this percentage over time. These two critical assumptions will be reviewed in future projection.

Interstate migration trends

One of the most important determinants of the projected size of the veteran population in a given state is the interstate migration of veterans. Campbell[5] describes the alternative approaches to migration used by the U.S. Census to project age and gender specific state populations to 2020. He indicates that a multi-state interstate migration projection overcomes many of the limitations of a net migration approach. Specifically, it eliminates the need for a raking procedure to assure that the total aggregated projected state populations equals the total national population. In the original VETPOP model, net migration estimates for civilians were applied to the veteran population. This imposes both the requirement of raking (to make sure state totals agree with the national totals) and the assumption that the veteran population has a similar net migration pattern as the civilian population.

There are two advantages of using civilian migration rates from the Census Bureau. First, the migration rates are based on a large number of observations (IRS administration records for twenty years on the entire U.S. population) and second, projected migration rates by age, gender, and race for twenty-five years into the future are readily available. However, Cowper[6] points out that the veteran population migrates at a higher rate than the civilian population[2]. Her work was based on comparing Census information for the 1970, 1980 and 1990 periods.

More recent data from the Current Population Survey when compared to VA internal data on veterans receiving compensation and pension benefits support many of Cowper's findings. Figure 3 provides information on the aggregate male civilian and male veteran interstate migration rates from the 1998 Current Population Survey and the 1999 rates for the C&P beneficiaries. If we apply the U.S. male civilian migration rates to the male veteran population (the second set of columns) we find that the overall age adjusted migration rate for veterans would be significantly lower (1.61% vs. the observed 2.11%). In contrast, if we apply the VA beneficiary-based interstate migration rates (the third set of columns) to the total male veterans population, the overall interstate migration rate is higher (2.59%). In fact a χ^2 goodness of fit tests reveal that the two candidate distributions of interstate migrations by age group (civilian migration and beneficiaries migration) are inadequate. Closer examination of the information in Figure 3 reveals that the differences lie in the higher age group. Currently, our tentative approach to incorporating migration rates in the new model is to apply veteran specific interstate migration rates for age groups less than 65, and, adopt the civilian interstate migration rates for the 65 and over age group.

Figure 4 provides the estimated out migration rates for three states using beneficiary data. This graph confirms the pattern of higher interstate migration rates for the younger age groups. For example, California veterans experienced an out migration of 5.3% of veterans in the age group 20-29 as compared to 1.3% for the 65 and over age group.

Projections and Output Tables

The new model will produce the same type of information as the current model concerning the total count of veterans, separations, deaths, and interstate migration, as well as, a variety of additional information concerning veteran characteristics. Two major types of reports will be produced from the new model:

- *Thirty- year national level reports* providing information on veteran counts for separations, and deaths by:
    - Period of service, gender and age (single year, five year age groups)

- Degree of disability, gender, and five year age groups
- Branch of service, gender and five year age groups
- Officer/Enlisted
- Gender.

2. Thirty-year state level reports providing information on veteran counts for separations, deaths and migration by:

- Period of service, gender and five year age groups.

While the model is run in the SAS programming language, the reports will be produced and distributed to users in Excel tables on CDs and through the VA web site. The use of pivot tables will permit users to develop additional output tables to meet special needs.

Summary

The new model is designed to build upon the experience gained by the VA in applying the original model. It incorporates enhancements to the original projection model based on studies documenting the improved mortality rates of veterans, greater understanding of their interstate migration patterns and the greater interchange of information between VA and the DoD. The model is PC-based and permits users to easily change the model's parameters and facilitate sensitivity analyses. Additional output reports are designed to serve the user community more effectively.

Bibliography

5. Campbell, Paul R. Population Projections for States by Age, Sex, Race and Hispanic Origin, October 1996, U.S. Census Bureau.

6. Cowper, Diane C. and Charles F. Longino, Jr. Veteran Interstate Migration and VA Health Service Use, The Gerontologist, Vol. 32, No. 1. 44-50.

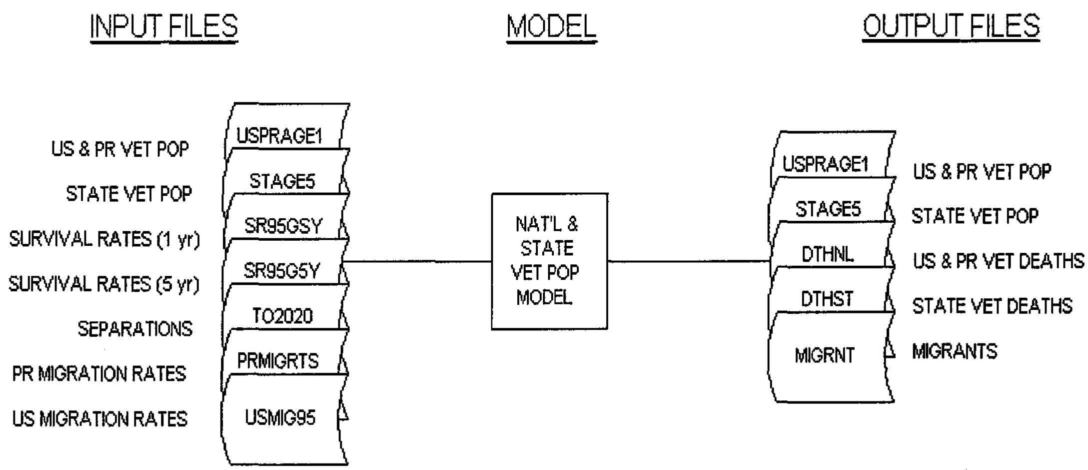
1. Department of Veterans Affairs, FY 1999 Annual Accountability Report, Statistical Appendix, 2000.

4. HayGroup, Barents Group, Mathematica Policy Research, Inc. The New Veteran Population Projection Model: Design and Specifications, Phase III.

2. Sorensen, Kathleen A. and Thomas C. Field, Projections of the U.S. Veteran Population: 1990 to 2010, Statistical Brief, Department of Veterans Affairs, May 1994.

3. Sorensen, Kathleen A., Technical Appendix, Projections of the United States and Puerto Rico Veteran Populations: 1990 to 2020.

FIGURE 1
INPUT AND OUTPUT FILES OF THE VETPOP MODEL



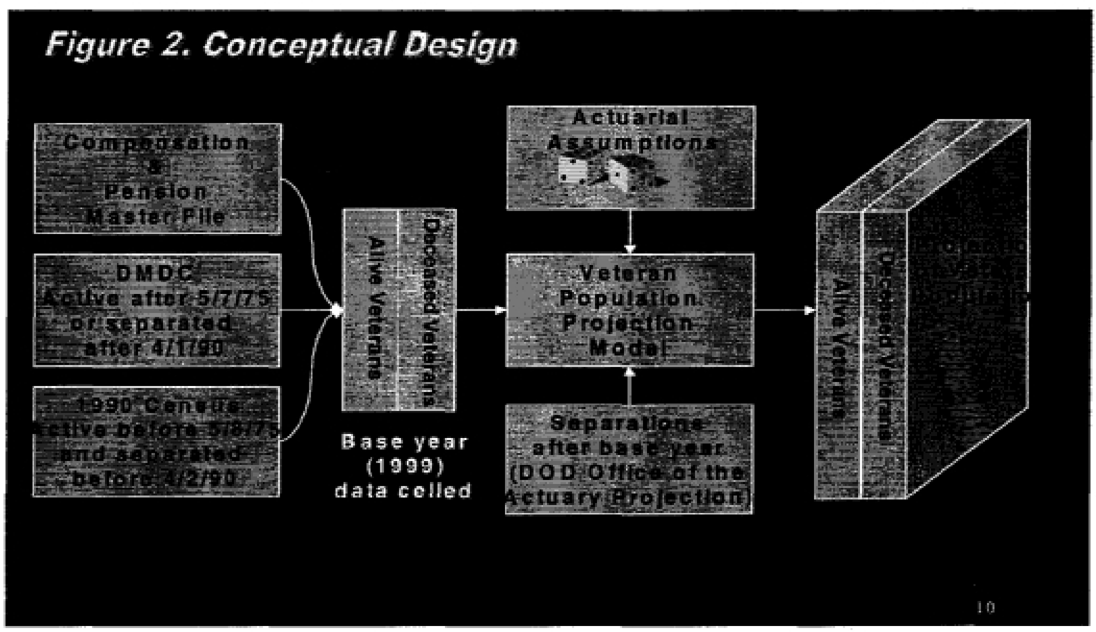
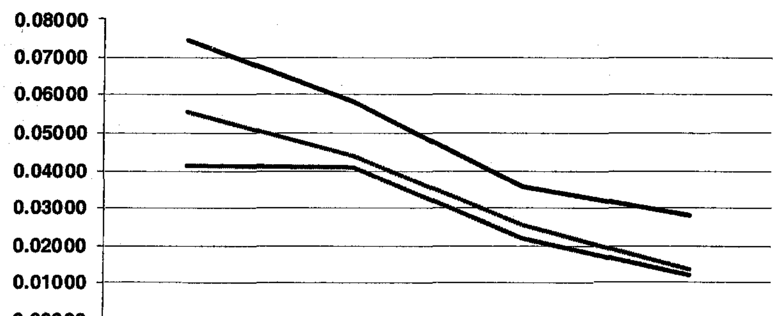


Figure 3. Differences in Migration Rate by Age

Age	Total Male US Pop-1998			Male All Veterans-1998			All C&P Beneficiaries- 1999		
	Total	Moved out of State	% Moved out of State	Total	Moved out of State	% Moved out of State	Total	Moved out of State	% Moved out of State
	(in '000)	(in '000)		(in '000)	(in '000)				
20 to 24 years	8826	385	4.36%	263	25	9.51%	11677	965	8.26%
25 to 29 years	9450	494	5.23%	787	62	7.88%	64126	4360	6.80%
30 to 44 years	32131	815	2.54%	3907	158	4.04%	382771	16991	4.44%
45 to 64 years	27272	446	1.64%	10331	193	1.87%	1106917	27466	2.48%
65 yrs. & over	13524	100	0.74%	8501	64	0.75%	1433136	18876	1.32%
Total	91208	2240	2.46%	23789	502	2.11%	2998627	66678	2.22%
Age Adjusted			1.61%			2.11%			2.59%
	1998 CPS March Supplement			1998 CPS Tables for Veterans (65-74 & 75+ COMBINED)			C&P Merged Files-3/2000-3/1999 (Veterans with no change in		

Figure 4. Out Migration Rates



	20-29	30-44	45-64	65+
Arizona	0.07447	0.05842	0.03559	0.02786
California	0.05537	0.04375	0.02532	0.01337
Connecticut	0.04134	0.04093	0.02158	0.01194

State

DYNAMIC PROGRAMMING OF FORECASTING APPARATUS

By Elliot Levy

Introduction of the DP Method

Dynamic Programming(DP) is a mathematical tool of Operations Research, a quantitative area of management science, for interrelated decision making.[1] There are **related decisions** involved in selecting a forecasting method, such as **type of *forecasting model*** and *the* **amount of independent variable input** which suggests the the DP technique to guide the forecaster in selection.

The DP approach, was introduced in the1950's by Richard Bellman, a pioneering systems engineer, at Rand Corporation, for **reducing the number of independent** variables as bottlenecks in stages of manufacturing.[2]

Prototype Application

The data applied in this example were from a horizon 1970-84 of annual forecasts of Commercial Building Construction, from an earlier paper of mine.[3] Standard percent of forecast error from seven types of models having up to five independent variables were extracted as input in this application.

The **standard percent of error** is the square root of squared forecast errors, % **Actual less Forecast**, divided by the horizon, as shown in the following formula. The **nf** is the number of forecasts that equal the horizon for computing an **average value** of the **percent of dispersion,** as shown in this formula.

$$S_f = \sqrt{\sum_{i=1}^{nf} (F-A)^2/nf}$$

An application of the above formula appears on the next page in Table I. In this particular example, 15 one year forecasts of a horizon of one were used to show how to compute the average percent error. Following Table I is another table, containing the percent errors of forecast by **type** of selected forecasting model by the **number** of variables, that had forecast results of greater error because these forecasts were made from a fifteen instead of a one-year time frame.

This computation example showed the following percent of forecast error:

$$Sf = \sqrt{\sum_{i=1}^{nf} (F-A)^2/nf} =$$

$$\sqrt{(1806.61 / 15} = \mathbf{8.8}$$

Table I contains the data for the above computation, after conversion to annual current dollars from the Appendix table. In both of the tables, errors of forecast were greater in the latest segment of the time period. However, their computed standard error was much smaller than that of the results from the models. Table I, with one year forecasts show this computation from the converted data.

The matrix containing **results over the fifteen-year** span is shown in **Table II.** These forecast errors here are **larger than those Table 1,** because forecast error is hypothetically larger when further away from past history. These models have structure similar to conventional econometric and time series forecasting methods, employing transformations of logarithms of and first differences from original data, **code chart** succeeding Table II.

TABLE I

% STANDARD ERROR OF FORECAST COMPUTATION,
ANNUAL U. S. COMMERCIAL BUILDING CONSTRUCTION,
YEARS 1970-84

(n) Years	(F) Forecast (Bil.$)	(A) Actual (Bil.$)	% (F-A) (Error)	% $(F-A)^2$
1970	10.8	9.8	10.2	104.12
1971	11.5	11.6	-0.9	0.74
1972	13.2	13.5	-2.2	4.94
1973	15.2	15.5	-1.9	3.75
1974	16.5	15.9	3.8	14.24
1975	12.0	12.8	-6.3	39.06
1976	12.3	12.8	-3.9	15.26
1977	14.0	14.8	-5.4	29.22
1978	18.2	18.6	-2.2	4.62
1979	23.3	24.9	-6.4	41.29
1980	31.0	26.6	16.4	268.26
1981	36.8	29.3	25.7	660.82
1982	38.8	34.2	13.6	184.22
1983	33.3	28.2	18.1	328.83
1984	35.4	32.0	10.4	107.23
				Sum=1806.61
				nf=15

Note: Actual and Forecast Converted to Current Dollars from 1972 and 1977 constant dollars.
Source: *US Industrial Outlook*. International Trade Administration. US Department of Commerce.

Table II

MATRIX OF % STANDARD ERRORS OF FORECAST [3]

Model # and #Variables	1	2	3	4	5	6	7
2	24	58	57	63	24	21	32
3	31	55	56	63	25	22	52
4	34	55	56	61	29	22	63
5	38	63	51	57	28	35	82

Linear Multiple Regression **Model** Codes:

1- Original Data:

2- First Absolute Differences ($X_t - X_{t-1}$)

3- First Percent Differences ($|(X_t - X_{t-1})| / X_t$)

4- Logarithms of Independent Variables **(X's)**: Growth Rate: **| $Y_t = f(\log X)_t$ |**

5- Logarithms of All Variables **(Y, X's)** : Elasticity: **| $\log Y_t = f(\log X)_t$ |**

6- Transfer Function (**Differences of Y on Differences of X; Moving Averages of Residuals of Actual data from fitted *function*)**

7- Distributed Lag [**Past Lagged Logarithmic Independent Variables (X's) Influence Upon Dependent Variable (Y)]**

Note: Both 6 & 7 are models that have **extended parametric length per X** as explanatory (**lagged** independent) variables of the Y (dependent)variable of interest.

Why Apply Dynamic Programming ?

The dynamic programming method was used for obtaining the optimal input needed from the models of this previous matrix from the accumulated minimal states of forecast error per model. The model types represent **stages** and the number of variables are the **states** in this problem.

Table III shows **ranking by forecast error,** which **does not show the same number of variables** at their lowest forecast error. The results show that there is no distinguishable solitary number of variables that would have minimized the forecast error. Therefore, a **dynamic programming** technique was needed to **solve for an optimal state.** No particular amount of input was evident as the optimal state, warranting a more **powerful** technique, **justifying** the application of **dynamic programming.**

TABLE III

INCONSISTENCY IN RANKING OF FORECAST ERROR

Rank of % Error Per Model by # Variables

	Model						
Variables	**1**	**2**	**3**	**4**	**5**	**6**	**7**
2	2	5	4	6	2	1	3
3	3	5	6	7	2	1	4
4	3	4	5	6	2	1	7
5	3	6	4	5	1	2	7

The DP Minimization Algorithm

Shown in this next example is the DP equation form of forecast error **minimization** by model.

Minimizing Error by Objective Function

$$\textit{Min}\ E(x)_{n+1} = \textit{Min}[\ F(x)_n + r(x)_{n+1}\ ],$$

where r is a remainder from the previous stage of a ***recursive*** process.

subject to the *behavior* of each ***state*** **(variables)** per ***stage*** **(models)**.

(Constraints or limits):

$1 \leq x_i \geq 4$, where {#Variables states}

$i = 1, ..., 7$ {Equation Type Stages}

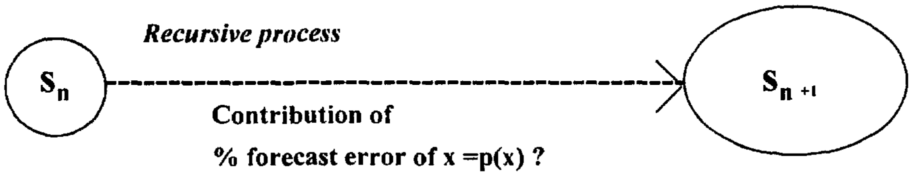
The task is to find, by an ***objective function,*** the minimum optimal state of error throughout succeeding stages. Richard Bellman used this same technique for optimum production in multi-stage manufacturing.[4] **This minimization equation** contained a remainder as the optimal solution from the previous stage, and when added to the next stage of forecast **results, made the succeeding computations cumulative.** Also, note that **actual forecast errors were applied** instead those derived from probability distributions. This is ***deterministic dynamic programming*** , **DDP,** where the forecast errors cannot be projected as estimates from distributions of probable error. Probable error relates to *stochastic dynamic programming.*

The computation ***process*** was ***recursive,*** having commenced with the last stage **n=7** until the first stage, **n=1**.
The **following diagram of DDP forecasting apparatus** showed the **recursive process** for this problem..

> **Note:** The following diagram and further computations throughout the remainder of this paper are from material borrowed from **Reference [1]**. Also, it was my idea to apply these analytical tools to forecasting and not anyone else in the establishment where I am employed. And they are exempt from an criticism relating to this paper.

DDP* DIAGRAM OF FORECASTING APPARATUS

Stage | **Stage**
n (n = model #7) | **n + 1 (n = model #6)**



% forecast error of x =p(x) ?
x=state

Accumulated Forecast Error states from consecutive % forecast errors through the last stage(s)

$f(s,xn)= p_n(x_n) + f^*_{n+1}(x_n)$

Accumulated Optimum with $p_n(x_n)$ contribution going toward minimization of the next past stage

$f^*_{n+1}(x_{n+1})$

***DDP** = *Deterministic Dynamic Programming*

The diagram of this DDP problem depicted going from a **current** in stage(n) into **an** optimal state for the **next** stage(n + 1) [5] that illustrated the minimum forecast error objective. The process began with the last stage **(n=7)** of the **Distributed Lag** forecasting equation and continued to the first stage (**n=1**) as the **Ordinary Least Squares** regression in simple or multiple form. Thus, the process in **reverse** order, similar to that of a door to door salesman deriving the best route to take from **previous sales *en route*.**

In the following computations, the **DDP method** was applied to accumulated percent of forecast error per stage for extracting the optimum(**lowest**) error and this optimum remainder became input for arriving at another optimum in the next recurrent stage. Decimal places were avoided by applying percent in whole numbers per iteration which prevented their vanishing into lost information.

In this segment, the following aspects are presented (1) the **states** by stage, (2) **data arrangement** by stage, and (3) **their formulas.**

1. The **states** were the **number of variables** from bi-variate to a Multi-variate model of five inputs. In the following **seven model stages, optimum states of input were derived for the forecasting models previously coded in Table II.**

2. The DDP computation procedure, started with a percent vector of de facto minimums indicated by f^*_7, and these values were used to develop square matrices in columns of forecast error in order to derive a second set of minimums. This process had ensued until one maximum value was obtained from a vector of the last column of minimums combined with the first column of the original data matrix, which adhered to the minimization algorithm already mentioned .

As shown in the next list, the process started with the Distributed Lag *(Code# 7)* forecast errors that were only a column of accepted minimums, and a subsequent matrix that incorporated these minimums, and finished with the final vector of the cumulative percent from applying the Ordinary Least Squares-Original Data*(Code# 1)*, bi or multi-variate.

Specifically by iterative stages , the process started with the Distributed Lag *column vector* of forecast errors**(Stage = 7)** and terminated recursively **(backwards)** to the Ordinary Least Squares-Original Data *row vector*(**Stage =1**).

Stage Listing of DDP Data for this problem

Stages by Model Type	Data Form
7	Vector (Column)
6	Matrix
5	Matrix
4	Matrix
3	Matrix
2	Matrix
1	Vector (Row)

3. The formulas for this problem have three components:

A. $p_n (x_n)$ = Percent of Forecast Error by column
B. $f^*_{n+1}(x_n)$ = Previous Column of Minimum Values
C. $f(s, x_n)$ = Cumulative sum of A & B

Thus, the DDP equation was: $f(s, x_n) = p_n (x_n) + f^*_{n+1}(x_n)$

Their notation per stage were:

DDP Formulas by Stage of Model Type

Model No.	Formula	Stage
7	$f(s, x_7) = f^*_7 (x_7)$	n=7
6	$f(s, x_6) = p_6 (x_6) + f^*_7(x_6)$	n=6
5	$f(s, x_5) = p_5 (x_5) + f^*_6(x_5)$	n=5
4	$f(s, x_4) = p_4 (x_4) + f^*_5(x_4)$	n=4
3	$f(s, x_3) = p_3 (x_3) + f^*_4(x_3)$	n=3
2	$f(s, x_2) = p_2 (x_2) + f^*_3(x_2)$	n=2
1	$f(s, x_1) = p_1 (x_1) + f^*_2(x_1)$	n=1

These recursive stage formulas represent calculated cumulative arrays of forecast error.

Solution

Mathematically, the basis of the solution was an objective function equation for derivation of the minimum state of forecast error for every state of Model type stage, as follows:

Minimize $\sum_{i=1}^{7} p_i(x_i)$ **(7 Stages of Forecast Error)**

Subject to: $\sum_{i=1}^{7} x_i = 4$ **(4 states per stage)**

For assessing the contribution of the past to current minimum in each state per stage shown in DDP diagram.

The following iterations for some stages of this DDP problem have been shown to demonstrate computation. **Those not shown are available from the author.**

1. Commencing stage **n=7**, as the first step, where each forecast error was **optimal**, because as the start of the iteration, **the last column of forecast error matrix were applicable only to this stage.**

n=7 **Vector of Forecast Errors: f_n*(s)**

s	f^*_7(s)	x^*_7
1	32	1
2	52	2
3	63	3
4	82	4.

2. The values for f^*_7 **(s)** ,optimal states of stage 7were in the last column of the data matrix in Table II, used in the computation of the first **matrix** by applying them to the sixth column of forecast error in Table II to derive the **next optimum** values, **f^*_6(s)**, of stage **n=6**, presented on the next page, in the **forecast errors of the Transfer Function model** in Stage 6. The first **f^*_7(s)** was added as a constant to all of the **$p_6(x_6)$** forecast errors and then the subsequent **f^*_7** optimums were applied to the **p_6** forecast errors by the same process in tandem.

3. These optimal f^*_6 **(s)** values **(s=1,..,4)** were the remainders of the previous stage added to original data of $p_5(x_5)$ of the fifth column in Table II, the first f^*_6 constantly added to all **$p_5(x_5)$** for column 1 of the next matrix of **stage n=5**. Also, the second **f^*_6** was applied to all of $p_5(x_5)$ data. These remaining optimal values, in tandem, were applied in the same process to complete **stage n=5**. From this matrix, optimal **minimum forecast errors** were extracted as **f^*_5 (s)** per x^*_5 for this 5th stage, in order to derive the minimum optimal values for the **forecast errors of the Elasticity model** of all data in logarithms.

STAGE 6: TRANSFER FUNCTION MODEL FORECAST ERRORS

n=6 **Matrix of Forecast Errors: $f^*_{n-1}(s)$**

$$f(s,x_6) = p_6(x_6) + f^*_7(x_6)$$

$f^*_7(s)$	$p_6(x_6)$	s	1	2	3	4	$f^*_6(s)$	x^*_6
32	21	1	53	73	84	103	53	1
52	22	2	54	74	85	104	54	1
63	22	3	54	74	85	104	54	1
82	35	4	67	87	98	117	67	1

STAGE 5: ELASTICITY MODEL FORECAST ERRORS

n=5 **Matrix of Forecast Error: $f^*_{n-2}(s)$**

$$f(s,x_5) = p_5(x_5) + f^*_6(x_5)$$

$f^*_6(s)$	$p_5(x_5)$	s	1	2	3	4	$f^*_5(s)$	x^*_5
53	24	1	77	78	78	91	77	1
54	25	2	78	79	79	92	78	1
54	29	3	82	83	83	96	82	1
67	28	4	81	82	82	95	81	1

4. In the following and final computation, **stage n=1**, a vector, contained the final minimum value of 267, from row $f(s, x_1)$, equal to optimum **$f^*_2(s)$** plus **$p_1(x_1)$**, the first column of Table II, **forecast errors from the Original Data model.**

Optimal Results

These **(f^*,x^*)** minimum values, **from stage 1 to 7**, were applied to derive the **optimal state of each stage** of forecasting model from the *s* column of each representative matrix and vector. These **optimal states per stage** were summarized in **Table IV**.

In each state s there was a **Minimum** value derived from each stage equation for x*, **the optimal state.** For example in **n=1**, the minimum value was **267**, the first cell of the vector of this stage of the **optimal state** x^*_1 used in tandem to pick the next optimal value of **243** at x^*_4 etc.

STAGE 1: ORIGINAL DATA MODEL FORECAST ERRORS

n=1	1	2	3	4
$f^*_2(x_1)$	243	240	240	240
$p_1(x_1)$	24	31	34	38
$f(s, x_1)$	**267**	**271**	**274**	**286**

$$f(s,x_1) = p_1(x_1) + f^*_2(x_1)$$

s \ x_1	1	2	3	4		$f^*_1(s)$	x^*_1
4	267	271	274	286		267	1

TABLE IV
OPTIMAL STATES FROM SOLUTION

Stage

n	s	x^*_n	Minimum	#Variables	Model No.	Description
1	4	1	267	2	1	Linear Regression: $Y_i=f(X_i)$
2	1	4	243	5	2	First (Absolute) Differences
3	4	4	185	5	3	First (Percent) Differences
4	4	1	134	2	4	Semi-logarithmic: log $Y_i=f(X_i)$
5	1	1	77	2	5	Logarithmic: Log Y_i=f(Log X_i)
6	1	1	53	2	6	Box-Jenkins Transfer Function
7	1	1	32	2	7	Polynomial Distributed Lag

Of these seven stages, a majority (5/7=71%) had minimum forecast error from using two forecast variables, one dependent and the other, a **predictor**.

x^*_n	%
2	71
5	29
7	100

A summarizing equation was made from these optimal results, $7x_n = 5x_1 + 0x_2 + 0x_3 + 2x_4$, which had extremes of smallest and the largest number of variables as forecasts entities, with **smallest as the major optimal state**.

Summary and Conclusions

The purpose Dynamic Programming(**DP**)was solving for the best **solution as a policy in guiding a forecaster toward *minimum* input usage in an equation for prediction** . *In this presentation, seven* types of forecasting models were observed. Their sporadic results were observed by their inconsistent ranking by minimal forecast error. In order to **confirm consistency**, the DP method was applied which corroborated **bi-variate** instead of multi-variate **relationships for a majority of the models.**

The DP algorithm solved for optimal minimums of forecast error from each state of model **forecast results, recursively,** by subsequent stage of model type to determine the best state of input usage per stage, in reverse order, beginning with the most complex and ending with the simplest model stage. In this problem, seven stages of **actual** forecast results were applied as input states having one to four predictors in various forecasting models for a ***Deterministic*** Dynamic Programming solution.

In this solution, seventy percent of the stages had **bivariate as the optimal outcome,** which is a ***clue*** *that* ***few*** *in* lieu *of* many predictors were sufficient for the most efficient forecasting equation.

REFERENCES

[1] Frederick S. Hillier and Gerald J. Lieberman. *Introduction to Operations Research.* (Third Edition) (Oakland, California: Holden-Day, Inc., 1980), p. 266.

[2] Bellman, Richard. *Dynamic Programming.* (Princeton, New Jersey: Princeton University Press, 1957), p.183.

[3] Elliot Levy, " Detecting Extraneous Forecast Variables by a Limiting Function, "*Proceedings of the Business and Economic Statistics Section.* (Alexandria, Virginia: American Statistical Association, 1988), p. 583.

[4] Richard Bellman, "*Some Problems in the Theory of Dynamic Programming,* " Econometrica (Volume22,1954), p. 38.

[5] *Hillier and Lieberman,* p. 271.

APPENDIX TABLE: COMPUTATION OF % STANDARD ERROR OF FORECAST FOR COMMERCIAL BUILDINGS CONSTRUCTION, 1970-84(BILLION$)

	ACTUAL REAL$ INPUT CONVERTED TO CUR$			FORECAST REAL$ INPUT CONVERTED TO CUR$			CONSTRUCTION COST INDEX						
YEAR	Current $	1972$	1977$	Current $	1972$	1977$	1972=100	1977=100	YEAR	ACTUAL (A)	FORECAST (F)	%(F-A)	$\%(F\text{-}A)^2$
1970	9.8			10.8					1970	9.8	10.8	10.2	104.12
1971	11.6			11.5					1971	11.6	11.5	-0.9	0.74
1972	13.5	14.6	19.7	13.2			100.0	64.1	1972	13.5	13.2	-2.2	4.84
1973	15.5			15.2					1973	15.5	15.2	-1.9	3.75
1974	15.9			16.5					1974	15.9	16.5	3.8	14.24
1975	12.8			12.0					1975	12.8	12.0	-6.3	39.06
1976	12.8			12.3					1976	12.8	12.3	-3.9	15.26
1977	14.8	13.0	14.8	14.0			156.0	100.3	1977	14.8	14.0	-5.4	29.22
1978	18.6	11.2		18.2			176.3	113.0	1978	18.6	18.2	-2.2	4.62
1979	24.9	12.9	19.0	23.3			200.2	128.8	1979	24.9	23.3	-6.4	41.29
1980	26.6	13.8	20.2	31.0			223.5		1980	26.6	31.0	16.4	268.26
1981	29.3		22.2	36.8	14.8		240.0	151.9	1981	29.3	36.8	25.7	660.82
1982	34.2		25.9	38.8	15.6	24.9		154.1	1982	34.2	38.8	13.6	184.22
1983	28.2		21.4	33.3		21.4		157.3	1983	28.2	33.3	18.1	328.83
1984*	32.0		24.3	35.4		24.9		163.7	1984	32.0	35.4	10.4	107.23

Note: Published Data originally in Real(1972 and 1977)dollars was converted to an annual trend of the above inputs in Current $.

Actual Converted to Current $:

1980: 1980(72$)/1979(72$)X1979(Cur$)=26.6

1981: 1981(77$)/1980(77$)X1980(Cur$)=29.3

1982: 1982(77$)/1981(77$)X1981(Cur$)=34.2

1983: 1983(77$)/1982(77$)X1982(Cur$)=28.2

1984: 1984(77$)/1983(77$)X1983(Cur$)=32.0

*Big jump in Other Commercial Construction

Forecast Converted to Current $:

1982: 1982(72$)/1981(72$)X1981(Cur$)=38.8

1983: 1983(77$)/1982(72$)X1982(Cur$)=33.3

1984: 1984(77$)/1983(72$)X1983(Cur$)=35.4

Sum $(F\text{-}A)^2$: 1806.61

Average$(F\text{-}A)^2$: 77.22

Standard Error of Forecast: $\boxed{8.8}$

(Square Root of Above Average)

Source: New Construction Put in Place: Trends and Projections(Years 1970-80) and Value of New Construction Put in Place(Years 1981-4). US Industrial Outlook, Years 1979-86.

FORECASTS, SOFTWARE, AND THE INTERNET

Chair: Stuart Bernstein
Bureau of Health Professions, U.S. Department of Health and Human Services

Seasonal Adjustment Using the X12 Procedure,
Tammy Jackson and Michael Leonard, SAS Institute, Inc.

Experiences With Placing ERS Food CPI and Expenditure Forecasts on the Web—Abstract,
Annette Clauson, Economic Research Service, U.S. Department of Agriculture

The DataWeb and DataFerrett: Accessing Data via the Internet—Abstract,
Bill Hazard, U.S. Census Bureau, U.S. Department of Commerce

SEASONAL ADJUSTMENT USING THE X12 PROCEDURE

Tammy Jackson and Michael Leonard
SAS Institute, Inc.

Introduction

The U.S. Census Bureau has developed a new seasonal adjustment/decomposition algorithm called X-12-ARIMA that greatly enhances the old X-11 algorithm. The X-12-ARIMA method modifies the X-11 variant of Census Method II by J. Shiskin A.H. Young and J.C. Musgrave of February 1967 and the X-11-ARIMA program based on the methodological research developed by Estela Bee Dagum, Chief of the Seasonal Adjustment and Time Series Staff of Statistics Canada, September 1979. The X12 procedure is a new addition to SAS/ETS software that implements the X-12-ARIMA algorithm developed by the U.S. Census Bureau (Census X12). With the help of employees of the Census Bureau, SAS employees have incorporated the Census X12 algorithm into the SAS System. The X12 procedure was experimentally introduced in Release 8.0, and after careful testing it was introduced for production in Release 8.1. It has since been enhanced for Release 8.2.

There have been numerous papers on the X-12-ARIMA algorithm. This paper provides a brief summary of the algorithm with references for the interested reader. It also summarizes the benefits of using the SAS System for Census X-12 seasonal adjustment/decomposition, briefly describes how to use the X12 procedure, and provides examples that compare the Census X-12 program to the X12 procedure. More details of the X12 procedure can be found in the *SAS/ETS Users Guide, Release 8.1.*

The X12 Procedure Summary

The X12 procedure seasonally adjusts monthly or quarterly time series. The procedure makes additive or multiplicative adjustments and creates an output data set containing the adjusted time series and intermediate calculations.

The X-12-ARIMA program combines the capabilities of the X-11 program (Shiskin, Young, and Musgrave 1967), the X-11-ARIMA/88 program (Dagum 1988), and introduces some new features (Findley et al. 1988). Thus, the X-12-ARIMA program contains methods developed by both the U.S. Census Bureau and Statistics Canada. The four major components of the X-12-ARIMA program are regARIMA modeling, model diagnostics, seasonal adjustment using enhanced X-11 methodology, and post-adjustment diagnostics. Statistics Canada's X-11 method fits an ARIMA model to the original series, then uses the model forecast and extends the original series. This extended series is then seasonally adjusted by the standard X-11 seasonal adjustment method. The extension of the series improves the estimation of the seasonal factors and reduces revisions to the seasonally adjusted series as new data become available.

Seasonal adjustment of a series is based on the assumption that seasonal fluctuations can be measured in the original series (O_t, t = 1,..., n) and separated from the trend cycle, trading-day, and irregular fluctuations. The seasonal component of this time series, S_t, is defined as the intrayear variation that is repeated constantly or in an evolving fashion from year to year. The trend cycle component, C_t, measures variation due to the long-term trend, the business cycle, and other long-term cyclical factors. The trading-day component, D_t, is the variation attributed to the composition of the calendar. The irregular component, I_t, is the residual variation. Many economic time series are related in a multiplicative fashion ($O_t=S_tC_tD_tI_t$) and others are related in an additive fashion ($O_t=S_t + C_t + D_t + I_t$). A seasonally adjusted time series, C_tI_t or $C_t + I_t$, consists of only the trend cycle and irregular components.

Summary of Usage

The X12 syntax contains the following statements:

```
PROC X12 options;
        BY variables;
        ID variables;
        TRANSFORM options;
        ESTIMATE;
        IDENTIFY options;
        REGRESSION options;
        ARIMA options;
        X11 options;
        FORECAST options;
        VAR variables;
        OUTPUT options;
RUN;
```

The PROC X12 statements perform basically the same function as the Census Bureau's X-12-ARIMA specs. *Specs* or specifications are used in X-12-ARIMA to control the computations and output. The PROC X12 statement performs some of the same functions as the Series spec in the Census Bureau's X-12-ARIMA software. The TRANSFORM, ESTIMATE, IDENTIFY, REGRESSION, ARIMA, X11, and FORECAST statements are designed to perform the same functions as the corresponding X-12-ARIMA *specs*, although full compatibility is not yet available.

The online help, online documentation, and printed documentation describe the X12 procedure syntax in greater detail. The Census Bureau documentation *X-12-ARIMA Reference Manual* can also provide added insight about the functionality of these statements. Appendix A contains a cross-reference between the X12 procedure and the X-12-ARIMA syntax.

Summary of Benefits

The X12 procedure is seamlessly incorporated into the SAS system. As with other analytical tools provided by SAS, this incorporation provides the following benefits:

Data Storage
Data can be efficiently stored in SAS data sets or warehoused in SAS data warehouses. Once data is stored in the SAS System, the X12 procedure and other analytical procedures can be used to analyze the data.

Data Preparation
The SAS language (DATA Step) of Base SAS can be used to prepare generic data for analysis. The EXPAND procedure of SAS/ETS software can be used to prepare time series data for time series analysis, decomposition, adjustment, modeling, and forecasting.

Output Delivery System (ODS)
ODS allows the output of the SAS procedures to be directed to a variety of destinations. These destinations include HTML (Web pages), Listing (Output Window), Printer (Network Printer), Output (SAS Data Set), and others. ODS also allows the format of the output to be customized as desired. In particular, the output of the X12 procedure can be customized to create reports specific to the needs of the organization.

Graphics
SAS/GRAPH software is the information and presentation graphics component of the SAS System. High-quality graphics can be generated for time series data. In particular, seasonal decomposition/adjustment graphs can be created from the data sets created by the X12 procedure.

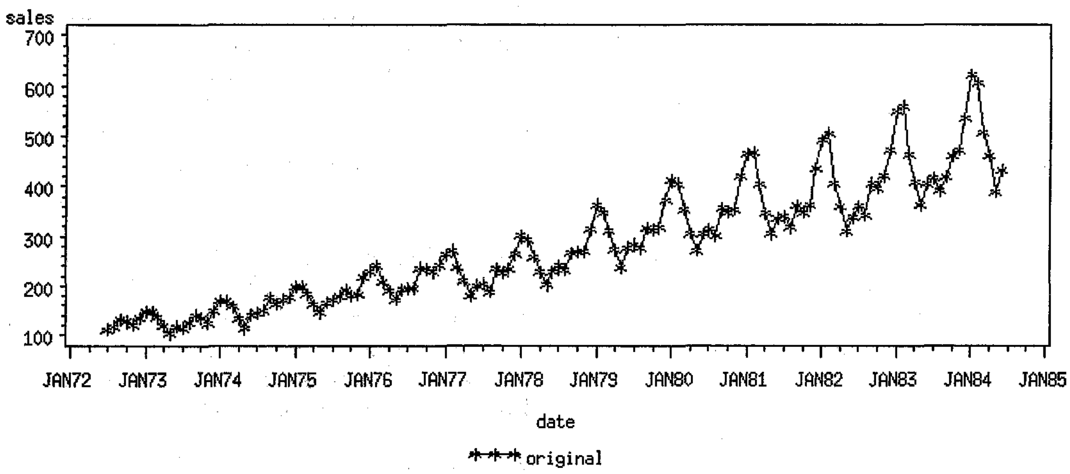
Application Development
SAS/AF (SCL based) or SAS/WebAF (Java based) applications can be custom-built for specific data analysis needs. In particular, applications for seasonal decomposition/adjustment using the X12 procedure and other analyses such as time series forecasting can be custom-built to address the specific needs of an organization.

Cross-Platform Compatibility
SAS programs and applications work on most major operating systems. SAS programs and applications developed on one platform can be used on other platforms

As shown, the SAS system provides many benefits for the seasonal decomposition/adjustment.

Examples of Usage

The following examples compare the syntax and output of the Census X-12 Spec File and the X12 procedure. Each of the following examples uses twelve years of monthly sales data (SALES). The sales data is plotted in the graph below.



Example 1

In this first example, the data is log transformed (POWER=0) and time series identification is specified. The IDENTIFY Spec in the Census X-12 program is compared to the IDENTIFY statement in the X12 procedure. As can be seen, the syntax is very similar. The IDENTIFY spec/statement determines the appropriate simple and seasonal differencing as well as tentatively identifying the ARMA(p,q)(P,Q)s orders.

EXAMPLE 1	
Census X-12 Spec File	**PROC X12 Code**
series{start=1972.07 data=(112 118 132 129 121 135 148 148 136 119 104 118 115 126 141 135 125 149 170 170 158 133 114 140 145 150 178 163 172 178 199 199 184 162 146 166 171 180 193 181 183 218 230 242 209 191 172 194 196 196 236 235 229 243 264 272 237 211 180 201 204 188 235 227 234 264 302 293 259 229 203 229 242 233 267 269 270 315 364 347 312 274 237 278 284 277 317 313 318 374 413 405 355 306 271 306 315 301 356 348 355 422 465 467 404 347 305 336 340 318 362 348 363 435 491 505 404 359 310 337 360 342 406 396 420 472 548 559 463 407 362 405 417 391 419 461 472 535 622 606 508 461 390 432)}	data sales; input sales @@; date = intnx('month', '01jul72'd, _n_-1); format date monyy.; datalines; 112 118 132 129 121 135 148 148 136 119 104 118 115 126 141 135 125 149 170 170 158 133 114 140 145 150 178 163 172 178 199 199 184 162 146 166 171 180 193 181 183 218 230 242 209 191 172 194 196 196 236 235 229 243 264 272 237 211 180 201 204 188 235 227 234 264 302 293 259 229 203 229 242 233 267 269 270 315 364 347 312 274 237 278 284 277 317 313 318 374 413 405 355 306 271 306 315 301 356 348 355 422 465 467 404 347 305 336 340 318 362 348 363 435 491 505 404 359 310 337 360 342 406 396 420 472 548 559 463 407 362 405 417 391 419 461 472 535 622 606 508 461 390 432 run;
	proc x12 data=sales seasons=12 date=date;
	var sales;
transform{power=0}	transform power=0;
identify{diff=(0, 1) sdiff = (0, 1)}	identify diff=(0,1) sdiff=(0,1);
	run;

Example 2

Continuing from the first example, the ARIMA Spec in the Census X-12 program is compared to the ARIMA statement in the X12 procedure. As can be seen, the syntax is similar. The ARIMA spec/statement specifies the simple and seasonal differencing as well as the ARMA(p,q)(P,Q)s orders.

EXAMPLE 2	
Census X-12 Spec File	**PROC X12 Code**
series{start=1972.07 data=(*...see datalines in example 1 ...*)}	data sales; input sales @@; date = intnx('month', '01jul72'd, _n_-1); format date monyy.; datalines; *...see datalines in example 1 ...* run;
	proc x12 data=sales seasons=12 date=date;
	var sales;
transform{power=0}	transform power=0;
arima {model=(0,1,1) (0,1,1)}	arima model=((0,1,1) (0,1,1));
estimate { }	estimate;
	run;

Example 3

Continuing from the second example, the X11 Spec in the Census X-12 program is compared to the X11 statement of the X12 procedure. The X11 spec/statement specifies X-11 decomposition.

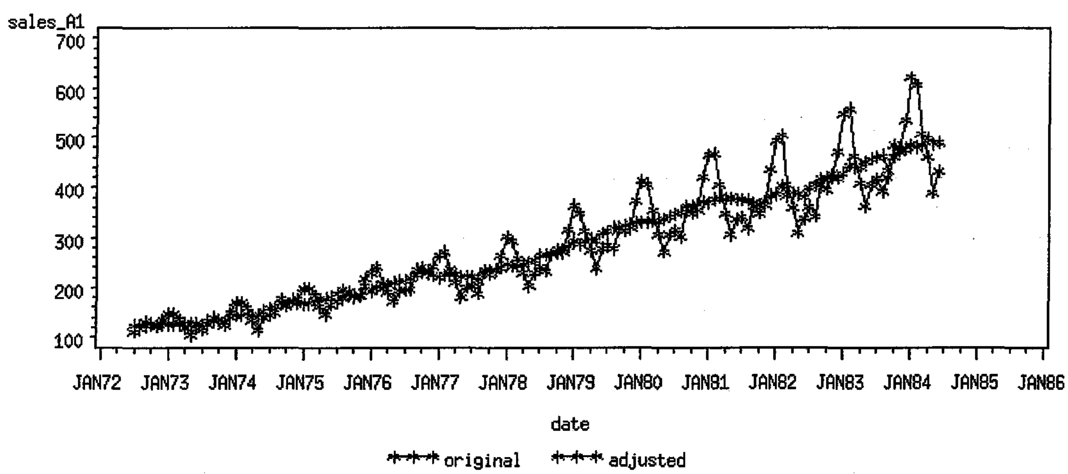
EXAMPLE 3	
Census X-12 Spec File	**PROC X12 Code**
series{start=1972.07 data=(*...see datalines in example 1 ...*)}	data sales; input sales @@; date = intnx('month', '01jul72'd, _n_-1); format date monyy.; datalines; *...see datalines in example 1 ...* run;
	proc x12 data=sales seasons=12 date=date;
	var sales;
transform{power=0}	transform power=0;
arima {model=(0,1,1) (0,1,1)}	arima model=((0,1,1) (0,1,1));
estimate { }	estimate;
x11{}	x11;
	run;

Example 4

Example 3 has been expanded to include an output statement. SAS/GRAPH is used to plot the original and seasonally adjusted series contained in the dataset.

EXAMPLE 4	
Census X-12 Spec File	**PROC X12 Code**
series{start=1972.07 data=(*...see datalines in example 1 ...*)}	data sales; input sales @@; date = intnx('month', '01jul72'd, _n_-1); format date monyy.; datalines; *...see datalines in example 1 ...* run;
	proc x12 data=sales seasons=12 date=date;
	var sales;
transform{power=0}	transform power=0;
arima {model=(0,1,1) (0,1,1)}	arima model=((0,1,1) (0,1,1));
estimate { }	estimate;
x11{}	x11;
	output out=out a1 d11;
	run;

	symbol1 i=join v='star';
	symbol2 i=join v='circle';
	legend1 label=none value=('original' 'adjusted');
	
	proc gplot data=out;
	plot sales_A1 * date = 1
	sales_D11 * date = 2 / overlay legend=legend1;
	run;
	quit;



Example 5

Here the results from Example 3 are directed to HTML files using the SAS Output Delivery System (ODS).

EXAMPLE 5	
Census X-12 Spec File	**PROC X12 Code**
series{start=1972.07 data=(*...see datalines in example 1 ...*)}	data sales; input sales @@; date = intnx('month', '01jul72'd, _n_-1); format date monyy.; datalines; *...see datalines in example 1 ...* run;
	Ods html file="out.html" contents="out_index.html" frame="out_frame.html";
	proc x12 data=sales seasons=12 date=date;
	var sales;
transform{power=0}	transform power=0;
arima {model=(0,1,1) (0,1,1)}	arima model=((0,1,1) (0,1,1));
estimate { }	estimate;
x11{}	x11;
	run;
	ods html close;

Exact ARMA Maximum Likelihood Estimation					
For variable sales					
Parameter	Lag	Estimate	Standard Error	t Value	Pr > |t|
Nonseasonal MA	1	0.40181	0.07887	5.09	<.0001
Seasonal MA	12	0.55695	0.07626	7.30	<.0001

Table F 2: Summary Measures					
Table F 2.F: Relative Contribution of the components to the stationary portion of the variance in the original series					
For variable sales					
I	C	S	P	TD&H	Total
0.39	11.27	87.04	0.00	0.00	98.70

Table F 3: Monitoring and Quality Assessment Statistics			
All the measures below are in the range from 0 to 3 with an acceptance region from 0 to 1.			
For variable sales			
1.	The relative contribution of the irregular over three months span (from Table F 2.B).	M1=	0.038
2.	The relative contribution of the irregular component to the stationary portion of the variance (from Table F 2.F).	M2=	0.039
3.	The amount of month to month change in the irregular component as compared to the amount of month to month change in the trend-cycle (from Table F2.H).	M3=	0.000
4.	The amount of autocorrelation in the irregular as described by the average duration of run (Table F 2.D).	M4=	0.875
5.	The number of months it takes the change in the trend-cycle to surpass the amount of change in the irregular (from Table F 2.E).	M5=	0.268
6.	The amount of year to year change in the irregular as compared to the amount of year to year change in the seasonal (from Table F 2.H).	M6=	0.700
7.	The amount of moving seasonality present relative to the amount of stable seasonality (from Table F 2.I).	M7=	0.198
8.	The size of the fluctuations in the seasonal component throughout the whole series.	M8=	0.435
9.	The average linear movement in the seasonal component throughout the whole series.	M9=	0.352
10.	Same as 8, calculated for recent years only.	M10=	0.461
11.	Same as 9, calculated for recent years only.	M11=	0.414

****** ACCEPTED *** at the level 0.26***
****** Q (without M2) = 0.29 ACCEPTED.***

Conclusion

The X12 procedure of SAS/ETS software is an adaptation of the U.S. Bureau of the Census X-12-ARIMA Seasonal Adjustment program. The X12 procedure is fully incorporated into the SAS system. This incorporation permits the storage and the preparation of data for subsequent analysis and for the presentation of the analysis using high-quality graphics, customized reports, and applications.

Acknowledgments

SAS Institute, Inc. is thankful for the support of the U.S. Census Bureau for the assistance provided in the development of the X12 procedure. In particular, Brian Monsell and Catherine Hood have contributed greatly. This paper and its associated practical demonstration relied heavily on the contributions of Evan Anderson, Virginia Clark, and Mark Traccarella of SAS Institute Inc.

References

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (1994), *Time Series Analysis: Forecasting and Control*, Third Edition, Englewood Cliffs, NJ: Prentice Hall, Inc.

Dagum, E. B. (1988), *The X-11-ARIMA/88 Seasonal Adjustment Method: Foundations and User's Manual*, Ottawa: Statistics Canada.

Findley, D. F., Monsell, B. C., Bell, W. R., Otto, M. C., and Chen, B. C. (1998), "New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program," *Journal of Business and Economic Statistics*, 16, 127 -176 (with Discussion).

Ladiray, D. and Quenneville B. (1999), "Understanding the X11 Method," Working Paper, Time Series Research and Analysis Centre, Statistics Canada.

Lothian, J. and Morry M., (1978), "A Test of Quality Control Statistics for the X-11-ARIMA Seasonal Adjustment Program," Research Paper, Seasonal Adjustment and Times Series Staff, Statistics Canada.

SAS Institute Inc. (2000), *SAS/ETS User's Guide, Release 8.1*, Cary, NC: SAS Institute Inc.

Shiskin, J., Young, A. H., and Musgrave, J. C. (1967), "The X-11 Variant of the Census Method II Seasonal Adjustment Program," Technical Paper No. 15, U.S. Department of Commerce, Bureau of the Census.

U.S. Bureau of the Census (1999), *X-12-ARIMA Reference Manual*, U.S. Department of Commerce, Washington, DC, [ftp://ftp.census.gov/pub/ts/x12a/].

U.S. Bureau of the Census (1999), "X-12-ARIMA Seasonal Adjustment Program," [ftp://ftp.census.gov/pub/ts/x12a/].

U.S. Bureau of the Census (1999), *X-12-ARIMA Quick Reference for Unix*, U.S. Department of Commerce, Washington, DC, [ftp://ftp.census.gov/pub/ts/x12a/].

Appendix A – Cross Reference of PROC X12 and X-12-ARIMA Syntax

SAS (V. 8.2) STATEMENT	SAS OPTION	DESCRIPTION	CENSUS SPEC	CENSUS ARGUMENT
PROC X12		Mostly data specifications	series{ }	
	DATA=	Should specify the input data set		data=
	DATE=	Date variable name		none equivalent
	START=	Date of 1st observation		start=
	SPAN=	(monyy,monyy) or ('yyQq','yyQq')		span=
	SEASONS=	4 for quarterly, 12 for monthly data		period=
	INTERVAL=	QTR or MONTH		period=
	NOPRINT	Suppress all printing	All specs	print=none
TRANSFORM		Transform or prior adjust series	transform{ }	
	POWER=	Box-Cox power transformation parameter		power=
	FUNCTION=	Transformation specified by name: NONE, LOG, SQRT, INVERSE, LOGISTIC, AUTO		function=
IDENTIFY		Used to identify the ARIMA portion of the model using seasonal and nonseasonal differencing	identify{ }	
	DIFF=	Orders of nonseasonal differencing		diff=
	SDIFF=	Orders of seasonal differencing		sdiff=
REGRESSION		reg information for regARIMA model	regression{ }	
	PREDEFINED=	List of predefined regression variables: CONSTANT, LOM, LOMSTOCK, LOQ, LPYEAR, SEASONAL, TD, TDNOLPYEAR, TD1COEF, TD1NOLPYEAR		variables=
ARIMA		ARIMA modeling information	arima{ }	
	MODEL= ((p d q)(P D Q)s)	Specify an ARIMA model (p d q)(P D Q)s using Box-Jenkins notation (if s is omitted, s=seasons)		model=
ESTIMATE		Estimates the regARIMA model specified by the regression and arima statements	estimate{ }	
X11		Seasonal adjustment info	x11{ }	
	MODE=	MULT, ADD, LOGADD, PSUEDOADD		mode=
	SEASONALMA=	Seasonal moving average used to estimate seasonal factors: S3X1, S3X3, S3X 5, S3X9, S3X15, STABLE, X11DEFAULT, MSR		seasonalma=
	TRENDMA=	Value for Henderson moving average		trendma=
	OUTFORECAST	Appends forecasts to tables A6, A8, A16, B1, D10, and D16		appendfcst=yes

FORECAST		Control forecast options	forecast{ }	
	LEAD=	The number of periods ahead to forecast		maxlead=
VAR		SAS standard statement to specify the time series variables to be adjusted/forecast		
BY		SAS standard statement to specify variables used in By-Group processing	none equivalent	
ID		SAS standard statement to specify variables used for identification purposes only		
OUTPUT		Information for output datasets for time series		
	out=	SAS-data-set name		
	A1	Original series	series{ }	save=(span)
	A6	regARIMA trading day component	regression{ }	save=(tradingday)
	A8	regARIMA combined outlier component	regression{ }	save=(outlier)
	B1	Prior adjusted or original series	x11{ }	save=(adjoriginal)
	C17	Final weight for irregular components	x11{ }	save=(irrwt)
	D8	Final unmodified S-I rations	x11{ }	save=(unmodsi)
	D9	Final replacement values for extreme S-I rations	x11{ }	save=(replacsi)
	D10	Final seasonal factors	x11{ }	save=(seasonal)
	D10D	Final seasonal difference	x11{ }	save=(seasonaldiff)
	D11	Final seasonally adjusted series	x11{ }	save=(seasadj)
	D12	Final trend cycle	x11{ }	save=(trend)
	D13	Final irregular series	x11{ }	save=(irregular)
	D16	Combined adjustment factors	x11{ }	save=(adjustfac)
	D16B	Final adjustment differences	x11{ }	save=(adjdiff)
	D18	Combined calendar adjustment factors	x11{ }	save=(calendar)
	E5	Percent changes in original series	x11{ }	save=(origchanges)
	E6	Percent changes in final seasonally adjusted series	x11{ }	save=(sachanges)
	E7	Differences in final trend cycle	x11{ }	save=(trendchanges)
	MV1	Original series adjusted for missing value regressors	series{ }	save=(missingvaladj)

Missing values are automatically imputed.

Experiences With Placing ERS Food CPI and Expenditure Forecasts on the Web

Annette Clauson
Economic Research Service, U.S. Department of Agriculture

Along with energy prices, food prices are the most volatile consumer price category that the U.S. government tracks. The only government entity that systematically examines food prices and provides food price forecasts is the Economic Research Service. As the forecaster of the Consumer Price Index (CPI) for several food categories, I developed a briefing room, Food Market Indicators, for the ERS web site three years ago. Along with the food CPI forecasts, this briefing room contains timely data on food expenditures, average retail food prices, food markets data, and food cost review data. Currently, this briefing room is the second most popular briefing room site on the ERS web site. In this session I will discuss my experiences of placing timely government forecasts and data on the Internet and the expectations of the customers and users of the data and information. I will also address our agency procedures for placing and posting forecasts on the ERS internet web site.

The DataWeb and DataFerrett: Accessing Data via the Internet

Bill Hazard
U.S. Census Bureau, U.S. Department of Commerce

The DataWeb is the infrastructure for intelligent browsing and accessing data across the Internet. The DataWeb brings together under one umbrella demographic, economic, environmental, health, and other datasets that are usually separated by geography and/or organization. The DataFerrett is the Browser for the DataWeb. DataFerrett, with its new Java 1.3 plug-in, accesses the data on the DataWeb and supports metadata searches across surveys, on-the-fly variable recoding, more complex tabulations, and graphics as well as other enhancements. Currently the DataFerrett provides access to data from the Current Population Survey and many of its supplements.

COMMODITY FORECASTS

Chair: Karen S. Hamrick
Economic Research Service, U.S. Department of Agriculture

Modeling Soybean Prices in a Changing Policy Environment,
Barry K. Goodwin, North Carolina State University
Randy Schnepf, Economic Research Service, U.S. Department of Agriculture
Erik Dohlman, Economic Research Service, U.S. Department of Agriculture

An Assessment of a "Futures Method" Model for Forecasting Season Average Farm Price for Soybeans,
Erik Dohlman, Economic Research Service, U.S. Department of Agriculture
Linwood Hoffman, Economic Research Service, U.S. Department of Agriculture
Randy Schnepf, Economic Research Service, U.S. Department of Agriculture
Mark Ash, Economic Research Service, U.S. Department of Agriculture

Cointegration Tests and Price Linkages in World Cotton Markets,
Stephen MacDonald, Economic Research Service, U.S. Department of Agriculture

Modeling Soybean Prices in a Changing Policy Environment

Barry K. Goodwin, North Carolina State University
Randy Schnepf, Economic Research Service, USDA
Erik Dohlman, Economic Research Service, USDA

Introduction

The oilseed products complex is an important component of the U.S. agricultural sector. In 2000, almost 75 million acres were planted to soybeans, representing over 29 percent of total planted acreage, making soybeans second only to corn in terms of acreage (ERS/USDA, 2000). Soybean acreage has increased steadily since 1990, when only 58 million acres were planted.

From a historical perspective, soybeans are rather unique in that they were not eligible for target-price deficiency payments nor were they subject to the explicit acreage restrictions of other program crops. However, the acreage-idling and base-acreage requirements, as well as government stock-holding behavior, of other program crops has indirectly affected soybean acreage decisions in the past.

Soybeans have been eligible for government price support loans for the past sixty years. In recent years, soybeans have benefited from a high loan rate relative to corn. This, coupled with eligibility for government marketing loan gains and loan deficiency payments, has stimulated production of soybeans.

Comprehension of the various factors underlying price determination is essential in order to understand the effects of policy changes and other shifts in market factors. Westcott and Hoffman (1999) considered the effects of market and policy factors using annual models of U.S. farm prices for corn and wheat. Their results confirmed the importance of the stocks-to-use ratio as an indicator of market supply and demand conditions. In addition, they used a number of discrete indicators of changing policy conditions. These indicators confirmed that changes in the policy environment can have important impacts on market prices and may influence the relationship between supply and demand factors and prices.

Such models have an important role in the development and validation of USDA projections of prices. Each month, the USDA analyzes major agricultural markets and publishes annual supply, demand, and price projections. Simple models relating price to observable supply and demand factors, such as the stocks-to-use ratio, are important tools in assessing predictions of such factors and price forecasts.

The objective of our analysis is to extend the models of Westcott and Hoffman (1999) by considering factors affecting U.S. soybean prices. We recognize that a more comprehensive specification of soybean price determination would incorporate the demand for soybean's joint products, meal and oil, in a larger multi-equation framework. But the goal of this research is to investigate the potential for using the simpler, single-equation stocks-to-use framework as an aid in monthly supply and demand analysis. Following Westcott and Hoffman (1999), we focus on the stocks-to-use ratio as an indicator of market supply and demand conditions. We also consider policy variables that may have impacted price relationships. Westcott and Hoffman (1999) focused on the 1975-1996 period. In contrast, we consider a much longer span of data and give explicit attention to the potential for structural changes in the relationships between prices and market factors.

We also focus on an issue not previously considered in evaluations of the relationship between the ending stocks-to-use ratio and prices—the potential endogeneity of these variables. One would certainly expect that prices adjust as supply is realized and as total use changes. However, demand theory suggests that total use will decline as prices increase—suggesting the potential for simultaneity between total use and prices. Even more likely, is the possibility that stock holding behavior is influenced by prices. Low prices typically serve as an incentive for agents to store a commodity in the hope that future market conditions will result in more favorable prices. Thus, ending stocks will be directly influenced by prices, making them endogenous in typical models relating prices to the stocks to use ratio.

The plan of our paper is as follows. The next section gives a brief review of factors suspected to be relevant to price determination in the U.S. soybean market. The third section presents an empirical analysis of price determinants for soybeans. We discuss structural change and endogeneity tests. In addition, we develop a gradual switching model that endogenizes the break point and speed of change inherent in the structural break. Improvements in the accuracy of model forecasts allowed by this parameter switching technique are identified and discussed. The final section of the paper includes a review of the analysis and offers some concluding remarks.

Conceptual Issues

Prices are determined by the interaction of supply and demand functions. Thus, a reduced-form expression for prices will relate prices to factors that influence supply and demand. As Westcott and Hoffman (1999) note, these factors are often summarized in the stocks-to-use ratio. Stocks adjust in response to shocks to supply and demand. Stocks will decrease in response to negative production shocks and will increase when production is high. Total use, which includes domestic consumption and exports, is generally more stable and tends to shift gradually over time. Of course, as we noted above, both factors may be simultaneously determined along with prices.

Following Westcott and Hoffman (1999) and Labys (1973), an equilibrium model for a storable commodity in a competitive market generally consists of a supply equation, a demand equation, a stocks equation, and an identity describing equilibrium. Supply (*S*) is a function of price (*p*) (or, more accurately, expected price) and factors (*z*) reflecting production shocks:

$$S_t = s(p_t, z_t). \quad (1)$$

Demand (*D*) is a function of prices and other demand shifters (*y*):

$$D_t = d(p_t, y_t). \quad (2)$$

Stocks (*K*) are influenced by prices and possibly other factors (*v*) reflecting storage costs and capacity constraints:

$$K_t = k(p_t, v_t). \quad (3)$$

Market equilibrium requires S_t - D_t - K_t = 0. This allows us to solve for a price-dependent reduced form expression that is a function of stocks and supply and demand shifters:

$$p_t = f(K_t, z_t, y_t). \quad (4)$$

Supply and demand shifters will include variables indicating changes in policy regimes as well as factors affecting weather and demand shocks. As noted above, it has become common to consider stocks in terms of the size relative to total usage. Thus, a common specification includes K_t/D_t, though, as we noted earlier in this paper, such a specification does not really represent a reduced form and thus may be subject to simultaneous equation biases. Further, to the extent that stock holdings are influenced by prices, K_t may also be endogenous to price.

In their analysis of corn and wheat prices, Westcott and Hoffman (1999) regressed prices (in logarithmic terms) on the logged ratio of total year-end stocks to use, the ratio of CCC held stocks to use, an interaction term that included a dummy variable representing the years 1978-85 and loan rate, and a dummy variable for 1986--- a year that was revealed to be an outlier in preliminary analyses. The years 1978-85 were singled out as a period when government intervention via the Famer-Owned Reserve (FOR) program, with high release prices and high loan rates relative to market prices, isolated significant amounts of corn and wheat from the market. Their wheat equation also included feed use and corn prices in the summer months, while excluding the 1986 dummy variable. Their empirical results confirmed a strong inverse relationship between the stocks to use ratio and price.

Empirical Analysis

We begin with a simple regression analysis of a form similar to that used by Westcott and Hoffman (1999) in their analysis of corn and wheat prices:

$$P_t = \forall_0 + \exists_1{*}(K_t/U_t) + \exists_2{*}\mathrm{LDP} + \exists_3{*}\mathrm{Drought} +$$

$$\exists_4{*}\mathrm{Loan\ Rate} + \exists_5{*}\mathrm{Loan\ Rate}{*}D_{78\text{-}85} \quad (5)$$

where all continuous variables are in logarithmic terms, LDP is a discrete indicator for the years in which significant loan deficiency payments were in made (1998 and 1999), Drought is a discrete indicator variable for drought years (1980, 1983, and 1988), and $D_{78\text{-}85}$ is a discrete indicator representing the period 1978-85. Westcott and Hoffman (1999) found that government programs had the most significant effect on prices during this period.

Data were collected from a variety of USDA sources. (An exact list of sources as well as the original estimation data are available from the authors on request.) The data span the period from 1942-1999. The soybean price is the season average price received by U.S. farmers. Stocks, denoted in Table 1 as Stocks_4, are ending stocks.

Estimates of the equation 5 (Model I) are presented in Table 1. Although the results suggest that this simple regression equation explains a considerable proportion of the variation in U.S. soybean prices, there are several reasons to question this specification. These concerns are related to structural shifts that may have occurred during the estimation period, the issue of price deflation, and endogeneity of stocks to use.

For example, one surprising result is that the overall stocks-to-use ratio does not appear to significantly influence soybean prices. The coefficient, though negative, is not statistically significant. For a shorter period of data (1975-1996), Westcott and Hoffman (1999) found a strong negative relationship between the stocks-to-use ratio and price, as would be expected. An examination of the data provides an explanation for this result.

Figure 1 illustrates the relationship between the stocks-to-use ratio and prices. A clear structural break in this relationship appears to have occurred around 1973. To the extent that this break is ignored, the estimates will suffer from specification biases.

Figure 1—Historical relationship between soybean price and stocks-to-use.

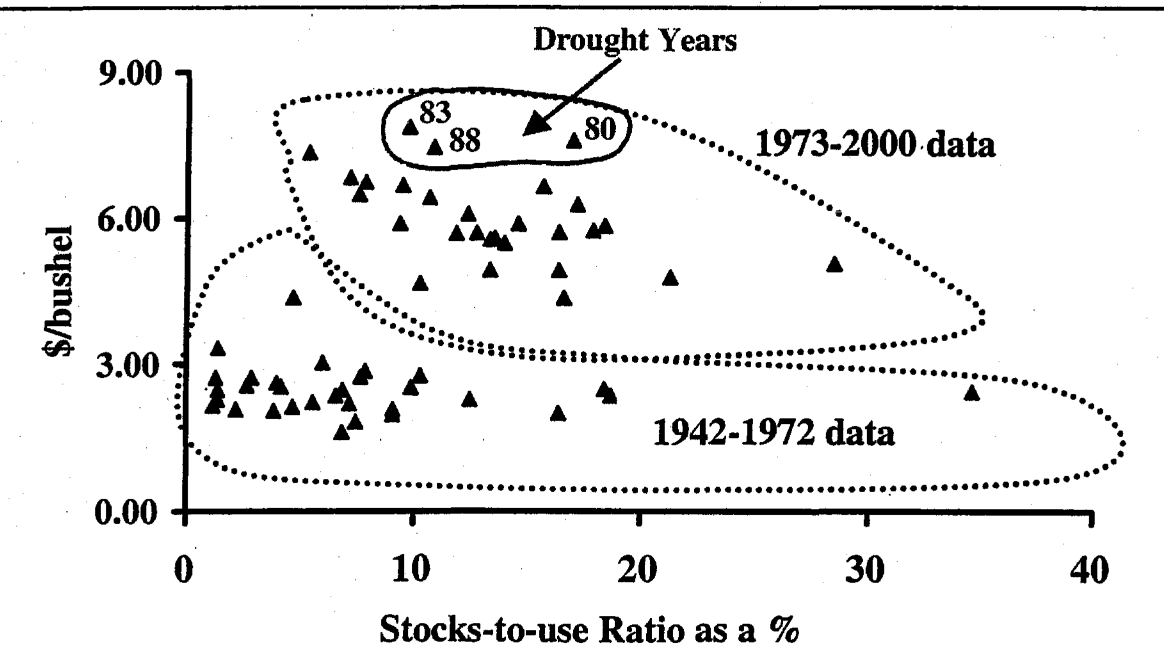


Table 1. OLS Estimates of Soybean Price Model

Variable	Model I		Model II		Model III	
Intercept	0.1458	(0.1195)	-3.0196	(0.5553)*	1.5416	(1.6075)
Drought	0.2195	(0.1620)	0.2639	(0.1268)*	0.0871	(0.1315)
(Loan Rate)*$D_{78\text{-}85}$	-0.0206	(0.0740)	0.0421	(0.0589)	0.0341	(0.0539)
LDP	-0.3338	(0.1938)*	-0.2515	(0.1522)*	-0.2921	(0.1411) *
Loan Rate	1.0718	(0.1124)*	-0.2295	(0.2385)	-0.0607	(0.2274)
$Stocks_4$/Use	-0.0102	(0.0494)	-0.0227	(0.0389)		
$Stocks_1/Use_{t-1}$					-0.6323	(0.2112)*
PPI_{t-1}			1.1250	(0.1935)*	0.9966	(0.1826)*
Adjusted R^2	0.7396		0.8353		0.8598	
Wu-Hausman Test	9.3025	[0.0037]*				
Chow Test at *1972/73*	*21.6300*	*[0.0001]**				

Note: $Stocks_1$ = 1st quarter stocks; $Stocks_4$ = ending stocks. Numbers in parentheses are standard errors. Numbers in brackets are probability values. Asterisks indicate statistical significance at the ∀ = 0.10 or smaller level.

A standard Chow test of the significance of this break was applied and found to be very significant, with an F-value of 21.6, which exceeds the critical values at all conventional levels of significance. We are unable to test for change in the drought, LDP, and loan rate–dummy variable interaction since these variables are all zero in the early (pre-1973) regime.

Another estimation issue involves the fact that nominal prices are the target of the analysis, and yet no adjustments are made for possible movements in the overall price level. The issue of deflating agricultural prices to account for movements in overall prices is a tricky one. It is widely recognized that real (i.e., deflated) agricultural prices have trended downward over time, although the general levels of nominal (non-deflated) prices have not changed significantly over time.

To account for inflation, we considered an alternative specification (Model II) that adds an indicator of the overall price level---the farm producer price index. The PPI was lagged one period to obviate any additional endogeneity concerns. This is of minor significance in light of its role as an indicator of long-run aggregate price movements.

This is a flexible alternative to actually deflating the prices since this specification nests a situation of actual deflation (implied by a coefficient value of 1) as well as any other adjustment that may be xmore suitable. The results would seem to suggest that the loan rate and the PPI are highly correlated. The loan rate loses its statistical significance in the new specification while the producer price index is significant with a value reasonably close to one. The in-sample explanatory power of the amended specification appears to be considerably higher than the simple specification.

Finally, in addition to possible mis-specification concerns regarding structural change and movements in aggregate prices, the aforementioned issues relating to the possible endogeneity of the stocks-to-use ratio are relevant to an evaluation of the simple specification. As we have noted, conceptual and intuitive considerations lead one to suspect that the ending stocks-to-use ratio may be jointly determined with prices. To evaluate this possibility, we consider standard Wu-Hausman tests of endogeneity. We assume that the ratio of the 1st-quarter stocks (December of the September-August crop year) to the preceding year's use (referred to as $Stocks_1/Use_{t-1}$ in Table 1) is exogenous to farm prices received. We use this as an instrument for ending stocks and conduct the Wu-Hausman test for endogeneity. The results are somewhat startling—the Wu-Hausman test strongly confirms the significance of endogeneity. The test statistic is 9.3, which exceeds the Chi-square critical value at conventional levels of significance. When the ending stocks-to-use ratio is replaced by this instrument (Model III), the stocks-to-use ratio reveals strong statistical significance and the expected negative effect on prices.

In summary, our results raise important concerns about the simple specification that uses ending stocks to use and ignores structural change. This is not to say that earlier papers (e.g., Westcott and Hoffman (1999)) necessarily ignored structural change. On the contrary, their focus on later periods of data for analysis reflects a recognition of the structural change issue. An analysis of shifts in the relationship between the stocks-to-use ratio and prices confirms a structural break that appears to have occurred in 1973. In addition, our intuition that the ending stocks-to-use ratio may be jointly determined with price is confirmed, suggesting the potential for biases in empirical results that ignore this issue.

A Switching Model of Structural Change

A variety of methods for modeling structural change have been proposed in the literature. Almost all such methods entail a shift or break in parameters over time. The simplest case involves the standard Chow test, in which a break at a predetermined point in the data is assumed. Of course, a problem associated with such an approach is that the timing of such a break must be known a priori. Alternatives to specifying the break prior to the test involve searching for the most significant break over a range of possible dates. Recent research by Andrews (1993) has demonstrated that conventional inference procedures are not applicable in such cases. In particular, the resulting F statistic is a *supremum* value over the range defined by the search space. The distribution of a *sup*(**F**) is not the same as a standard **F** and thus alternative inferential procedures are needed.

In addition to the issues associated with searching for a break point, conventional methods for modeling structural change are limited by the fact that they typically assume that such change occurs instantaneously. Although abrupt structural shifts are certainly possible, one would expect that gradual structural change is more likely to occur in economic relationships. Thus, a method which allows the data to choose the break point and the speed of adjustment between regimes is desirable. In this vein, we utilize a gradual switching regression method.

Gradual switching regressions were introduced by Tsurumi, Wago, and Ilmakunnas (1986). In contrast to their approach, we utilize a smooth transition function to represent the speed and timing of a structural shift between regimes. The use of transition functions as a means for modeling structural shifts was introduced by Bacon and Watts (1971). In our analysis, we allow the shift to occur gradually and identify the timing and speed of the shift using our estimation data. In particular, we represent structural change in terms of a shift in the parameter set from $\exists^{(I)}$ to $\exists^{(II)}$. A mixing term 8_t, that is constrained by construction to lie in the open interval (0,1), is used to represent shifting between regimes. Our specification of the mixing problem allows us to rewrite the simple regression relationship considered above y = X∃ as:

$$y_t = (1-8_t)\, X_t\, \exists^{(I)} + 8_t\, X_t\, \exists^{(II)} + e_t. \qquad (6)$$

The mixing term 8_t is given by:

$$8_t = M((t-:)/\Phi) \quad t = 1,\ldots,N; \qquad (7)$$

where M is the normal cumulative distribution function (cdf) and : and Φ are parameters to be estimated. Our smooth transition function approach has much in common with the smooth threshold modeling techniques of Terasvirta (1994). A similar approach to specification and estimation is undertaken there, though in that case observations may switch between regimes more than once. In our approach, the regime switch is permanent.

Note that : represents the observation lying one-half way between regimes 1 and 2 (i.e., for which $8_t = 0.50$). The bandwidth parameter Φ represents the speed of adjustment between regimes, with larger values of Φ corresponding to more gradual adjustments between regimes. Note that $\lim_{x\infty+4} M(x) = 1$ and $\lim_{x\infty-4} M(x) = 0$. (In reality, all observations fall between regimes given the asymptotic nature of the transition function, which never actually reaches zero from above or one from below.)

Estimation of the switching regression model may pose challenges. Though estimation follows standard nonlinear regression methods, identification issues may arise as the break point : nears either end of the data and as the speed of adjustment becomes very fast (i.e., as Φ approaches zero). We adopt the following estimation approach in this analysis. We first consider a standard grid search over possible values of : and Φ. We select the values that minimize a sum of squared error criterion (or, equivalently, that maximize an F-test of the specification against one without structural shifts). The optimal values of : and Φ are then used as starting values in a standard nonlinear regression model.

Estimates of the gradual switching regression models are presented in Table 2. Two alternative specifications are considered. The first includes only loan rates and the stocks-to-use ratio (using the ratio of 1^{st}-quarter ending stocks to last year's use). The second includes dummy variables representing drought years and the LDP as well as the producer price index. (Note that we do not allow the parameter on the producer price index to shift. Estimates of such a specification were numerically unstable.) In both cases, the : estimates for both models indicate a strong and immediate structural break centered at observation number 31, corresponding to 1972. Furthermore, the Φ estimates are quite large (0.89 in Model 1 and 0.87 in Model 2) suggesting a very rapid adjustment phase of approximately 2-3 years. Thus, the results are consistent with the Chow tests reported earlier as well as with earlier research that has argued in favor of structural breaks at this point in time. The speed and timing of the structural shift in the two single-equation models is illustrated in Figure 2.

The gradual switching model allows us to not only identify the timing and speed of structural shifts but also to characterize the nature of the shifts. In both models, the results suggest that the negative influence of the stocks-to-use ratio is much stronger in the latter period. In Model 1, the coefficient changes from -0.42 in the early regime to -0.70 in the latter regime. Likewise, in Model 2, the shift is from -0.42 to -0.61. The effect of loan rates on soybean prices also appears to vary from period to period. In the first regime, the coefficient on loan rates is statistically significant with a value of about 0.83. In the second regime, loan rates do not appear to have influenced prices. The addition of discrete indicators for drought and the LDP program and the inclusion of the producer price index as an indicator of general price movements do not appear to significantly alter these the results. When local market prices fall below the loan rate, the marketing loan program (LDP) allows producers to capture the price difference as a payment from the government. Prior to implementation of the marketing loan program, when market prices fell below the loan rate farmers would cede their crops to the government in return for the loan rate. Thus, the marketing loan program prevents the loan rate from acting as a floor for market prices. This negative effect on average market-prices is captured by the LDP variable.

Figure 2-- Estimated Transition Function for Single Equation Model.

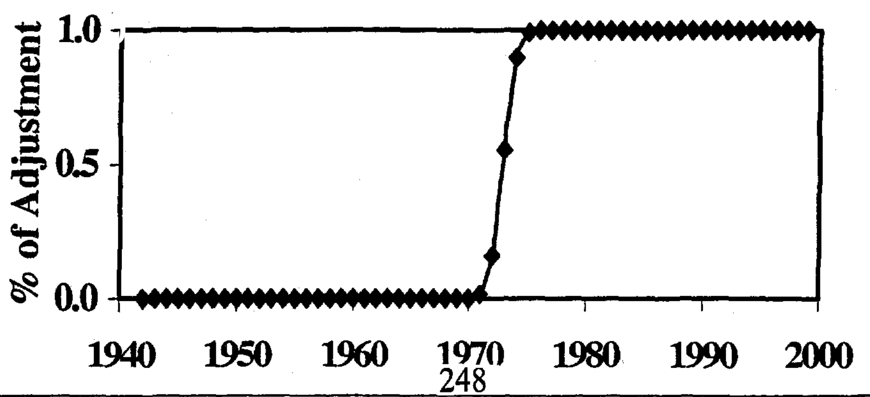


Table 2. Estimates of Gradual Switching Soybean Price Model

Variable	Model 1		Model 2	
:	30.8841	(0.2149)*	30.8562	(0.1933)*
Φ	0.8879	(0.2966)*	0.8678	(0.2679)*
Regime I: Intercept	3.0207	(1.0552)*	3.0677	(0.9966)*
Regime I: Loan Rate	0.8300	(0.1932)*	0.8390	(0.1939)*
Regime I: ($Stocks_1/Use_{t-1}$)	-0.4162	(0.1545)*	-0.4183	(0.1346)*
Regime II: Intercept	6.6061	(0.9069)*	5.9957	(1.1540)*
Regime II: Loan Rate	-0.0427	(0.0987)	-0.0027	(0.1120)
Regime II: ($Stocks_1/Use_{t-1}$)	-0.7042	(0.1380)*	-0.6137	(0.1585)*
Drought			0.0781	(0.701)
LDP			-0.2451	(0.0652)*
PPI_{t-1}			-0.0107	(0.1178)
Adjusted R^2	0.9598		0.9695	

Note: Regime I represents the pre-switching estimates; Regime II represents the post-switching estimates. $Stocks_1$ = 1st quarter ending stocks. Numbers in parentheses are standard errors. Numbers in brackets are probability values. Asterisks indicate statistical significance at the ∀ = 0.10 or smaller level.

In summary, the results are largely consistent with the findings of earlier research. A structural shift does indeed appear to have characterized market price relationships in the reduced form model of soybean farm prices. The shift appears to have occurred at about 1972-73 and appears to have been very rapid.

Concluding Remarks

An understanding of fundamental reduced form relationships among variables important to supply and demand and market prices is important to commodity and policy analysts. This paper reports on an analysis of such market relationships for soybeans. Following earlier research, we considered a simple regression model for annual soybean prices that included the stocks-to-use ratio, the loan rate, and a number of discrete indicators of policy. We pursue two distinct issues in our consideration of this relationship.

The first involves explicit modeling of structural change. A primary focus of our analysis involved the identification and characterization of structural shifts. We utilize models of discrete structural breaks as well as an alternative gradual switching regression approach that permits change to occur gradually. Our results confirm the significance of an abrupt structural break that occurred at about 1973-74. The timing and speed of the adjustment were robust over a number of alternative specifications. The results suggest that soybean prices have become more sensitive to relative stocks.

A second focus of our analysis involves the potential endogeneity of the stocks-to-use ratio and prices. Theoretical considerations of stockholding behavior suggest that stocks will be affected by prices. Likewise, total use should be negatively influenced by prices. We conduct explicit tests of this endogeneity and confirm that significant biases may arise if the endogeneity of the stocks-to-use ratio is ignored in a reduced form price equation.

The early 1970s was a period of significant changes in world agricultural markets when nearly two decades of fairly stable commodity prices ended with a dramatic spike. This tumultuous period was marked by an unexpected surge in world grain demand and trade, coupled with poor harvests and rapid, dynamic macroeconomic changes (Riley; 1996). An emergence of international markets from the post-Bretton Woods period enhanced international trade in agricultural commodities. In addition, significant development of soybean production in other competing (Southern Hemisphere) markets occurred during this period. Thus, it is not surprising that structural relationships for soybean prices appear to have shifted during this period.

Future research will consider the development of explicit tests for structural change in the gradual switching context. These tests are complicated by the widely recognized problem of a set of parameters that are unidentified under the null hypothesis of no structural change. A variety of tests have been developed for such cases by Hansen (1997). Subsequent work will involve the application of these tests to the results presented here.

References

Andrews, D. W. K. "Tests for Parameter Instability and Structural Change with Unknown Change Point," *Econometrica* 61(1993):821-56.

Bacon, D. and D. G. Watts. "Estimation of the Transition Between Two Intersecting Straight Lines," *Biometrika* 58(1971):525-34.

Economic Research Service, USDA. "U.S. Acreage Expands," *Agricultural Outlook* August, 2000, pages 2-3.

Goodwin, Barry K. "Forecasting in the Presence of Structural Change: An Application of Gradually Switching VAR Systems." *Southern Journal of Agricultural Economics*, 24(December 1992):11-22.

Hansen, B. E. "Inference in TAR Models," *Studies in Nonlinear Dynamics and Econometrics* 2(April 1997) (online).

Labys, W. C. *Dynamic Commodity Markets: Specification, Estimation, and Simulation*, Lexington, MA: Lexington Books, D.C. Heath and Co., 1973.

Riley, Peter. "Global Grain Markets in 1996: Shades of 1972-74?" *Agricultural Outlook* September, 1996, pages 2-6.

Terasvirta, T. "Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models," *Journal of the American Statistical Association*, 89(1994): 208-218.

Tsurumi, H., H. Wago, and P. Ilmakunnas. "Gradual Switching Multivariate Regression Models with Stochastic Cross-Equational Constraints and an Application to the KLEM Translog Production Model," *Journal of Econometrics* 31(1986): 235-53.

Westcott, P. C. and L. A. Hoffman. "Price Determination for Corn and Wheat," Economic Research Service Technical Bulletin No. 1878, Washington, D.C., July 1999.

AN ASSESSMENT OF A "FUTURES METHOD" MODEL FOR FORECASTING SEASON AVERAGE FARM PRICE FOR SOYBEANS

Erik Dohlman, Linwood Hoffman, Randall Schnepf, and Mark Ash, U.S. Department of Agriculture, Economic Research Service

Introduction

The U.S. Department of Agriculture (USDA), in its efforts to provide reliable market information on agricultural products, develops short-run forecasts of production, use, and trade for numerous agricultural commodities, including soybeans. Based on expected supply and demand conditions, USDA also issues forecasts of annual commodity prices on a monthly basis, and these projections are used as an important planning tool by both the private and public sectors. For producers, forecasts of season-average farm price (SAFP) can affect marketing decisions. Furthermore, producers and users of agricultural commodities rely on forecasts to manage income and price risk. For policy-makers, accurate forecasts can be important for budgetary purposes related to farm programs.

Given the importance of price forecasts to market participants, the objectives of this study are twofold. First, we construct an alternative set of monthly soybean season-average farm price forecasts using the "futures method" model previously developed by Hoffman and Davison (1992), and assess the accuracy of these forecasts by comparing them with actual season-average farm prices during crop years 1981/82 to 1998/99. Second, we compare the accuracy of futures method forecasts to those published monthly by USDA in the *World Agricultural Supply and Demand Estimates* (WASDE) report. Our aim is to determine whether the futures method represents a generally reliable approach to forecasting commodity prices, as well as to provide an overall assessment of WASDE and futures method forecast accuracy.

In addition to our main objectives, we also explore whether the accuracy of futures forecasts improves when futures markets gain access to new information from the most recent WASDE report. That is, are forecasts based on futures prices immediately following the release of WASDE more accurate than those made just prior to the WASDE release. Intuitively, this makes sense. WASDE SAFP projections represent the sum of all publicly available market-related information, but some of this information, such as USDA National Agricultural Statistics Service (NASS) survey-based data on crop yields, are not made available to the public until the WASDE's release. Although market participants may anticipate this information, futures forecasts following the release of the WASDE should represent the most up-to-date composite of public and privately held information. To test this conjecture, we develop two separate forecasts of SAFP using the futures method – one based on futures price data available prior to the release of WASDE, and the other based on futures price data immediately following the release of WASDE.

The following section describes the method used to develop monthly forecasts of annual season-average soybean prices with futures, and illustrates the method with a November 1999 forecast for the 1999/2000 crop year. We then compare the historical accuracy of the futures forecasts with WASDE forecasts by calculating the mean absolute percentage error (MAPE) of the forecasts during crop years 1981/82 to 1998/99. Next, the average (1981/82 to 1998/99) absolute percentage error for each forecast month is examined separately to see if there is any pattern to differences between the alternative forecasts over the course of the crop year. We conclude with a brief summary.

Overview of Futures Forecasting Method

Using the futures method, forecasts of monthly average prices received by U.S. farmers are made for each month of the crop year starting with September. Price forecasts for each month of the crop year are initially based on the current month's futures price for the nearest contract maturing after the month being forecast (referred to as the "nearby futures contract").

Most market participants understand that the futures market is a composite indicator of anticipated supplies and demands and that current futures prices therefore provide important information about cash prices on future dates. However, participants also need to be able to forecast a price at the location and time when they plan to buy or sell. Thus, they need to predict the "basis," the difference between the futures price and the local price.

The futures method employed here uses an historical monthly average basis (historical monthly farm price received minus historical monthly average futures price for the nearby contract) that is subtracted from the current nearby futures prices to yield a monthly U.S. average farm price forecast for each month of the crop year. The 12 monthly price forecasts are then multiplied by their five-year historic share of annual marketings and summed to produce a weighted season-average farm price forecast. As estimated monthly farm prices become available, the predicted season-average farm price becomes a composite of actual and forecasted prices.

Basis

The difference between a farm (henceforth "cash") price received at a specific location and the price of a particular futures contract is known as the basis. The basis tends to be more stable or predictable than either the farm price or futures price. Factors that can affect the basis include local supply and demand conditions for the commodity and its substitutes, handling costs, transportation and storage costs, and market expectations. The basis used in this analysis is a composite of these factors and represents an average of U.S. conditions.

The basis in this study is defined as the difference between the monthly U.S. average cash price received by producers and the monthly average settlement price for the nearby futures contract. For example, the September basis is the difference between the September average cash price received by producers less September's average settlement price of the November futures contract. A five-year moving average of these bases, used to eliminate distortions that may occur in any given year, is updated at the end of each crop year. Thus, data for the 1976 through 1980 crop years establish the historical basis used to develop the 1981 crop year futures forecast.

Data

Historical daily soybean futures settlement prices for crop years 1976 to 1999 are obtained from *TechTools* data service. Historical cash prices were acquired from USDA's (NASS) *Agricultural Prices*, and weights for monthly marketings were obtained from USDA's (NASS) December issues of *Crop Production* (prior to 1998) and November issues of *Agricultural Prices* (1998 to present).

Procedure and Illustration of futures method

Table 1 illustrates the method used to forecast the 1999/2000 crop year season-average soybean price in November 1999. Although the futures method forecast for 1999/2000 has been updated through August 2000, we present the November 1999 forecast to more clearly illustrate that SAFP forecasts are, in general, a composite of actual and forecasted monthly prices. It should be noted that our assessment of the accuracy of the futures method for crop years 1981/82 to 1998/99 is based on all twelve monthly forecasts for each year. Recall that we use the futures method to produce two alternative forecasts of the SAFP – one using a two-day average futures settlement price available just prior to the release of that month's WASDE, and one using a two-day average settlement price following the WASDE release. For simplicity of presentation, only the first (pre-WASDE) forecasts are shown in Table 1.

Seven steps are involved in the forecast process, illustrated here with the November 1999 forecast of the 1999/2000 crop year SAFP:

1. Futures settlement prices are gathered for the contracts that will mature during the forthcoming year (line 1). When pre-WASDE settlement prices are used, the two-day average futures price for the January, March, May, July, and September (2000) contracts available on November 8th and 9th were selected (WASDE was released on November 10). Estimates of actual monthly prices received are available from NASS and used for September and October 1999. The October 1999 price represents a mid-month estimate published in that month's issue of *Agricultural Prices* (the price is updated the following month). The November 1999 contract is not used for reasons discussed below.

2. The monthly futures prices are based on the settlement prices of the nearby contracts. For example, the futures prices for November and December represent the November (8th and 9th) average settlement price of the nearby January contract. The futures prices for January and February are based on the November settlement prices for the nearby contract for those months (March). During months in which a futures contract matures, the next contract month is used because futures contracts are affected by a decline in liquidity during the month of maturity. Although the September 2000 futures contract falls outside of the current crop year, this contract is used to establish the monthly futures price for August 2000.

3. A five-year moving average of the basis (cash prices minus the monthly average settlement price for the nearby futures contract) for each month is entered (on line 3).

Table 1 – Futures forecast of U.S. soybean season-average farm price, 1999/2000 crop year (November 1999)

Item	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep
					Dollars/Bushel								
1. Current futures price 1/ by contract					4.81		4.87		4.93		4.98	4.97	5.03
2. Monthly futures price based on nearby contract			4.81	4.81	4.87	4.87	4.93	4.93	4.98	4.98	4.97	5.03	
3. Plus the historical basis (cash less futures) 2/	-0.07	-0.25	-0.30	-0.23	-0.18	-0.19	-0.26	-0.26	-0.26	-0.20	-0.11	0.04	
4. Forecast of monthly average farm price			4.51	4.58	4.69	4.68	4.67	4.67	4.72	4.78	4.86	5.06	
5. Actual monthly farm price	4.57	4.49											
6. Spliced actual/forecast monthly farm price	4.57	4.49	4.51	4.58	4.69	4.68	4.67	4.67	4.72	4.78	4.86	5.06	
Annual price projection													
7. Marketing weights (percent)	6.9	22.8	9.2	7.4	13.6	7.2	7.4	5.6	4.7	4.8	5.4	5.1	
8. Weighted average forecast ($/bushel)			4.64										

1/ Contract months for soybeans include: September, November, January, March, May, July, and August.
2/ Data shown here are the 5-year average for crop years 1994-1998.

4. A forecast of the monthly average farm price (line 4) is computed by adding the basis (line 3) to the monthly futures prices (line 2), except when NASS monthly or mid-month price estimates are known.

5. The NASS monthly average farm price is entered on line 5 as it becomes available. In this example, the September price is for the entire month and the October price is a mid-month estimate. In December, the estimate for October would be updated and a mid-month estimate for November would be included.

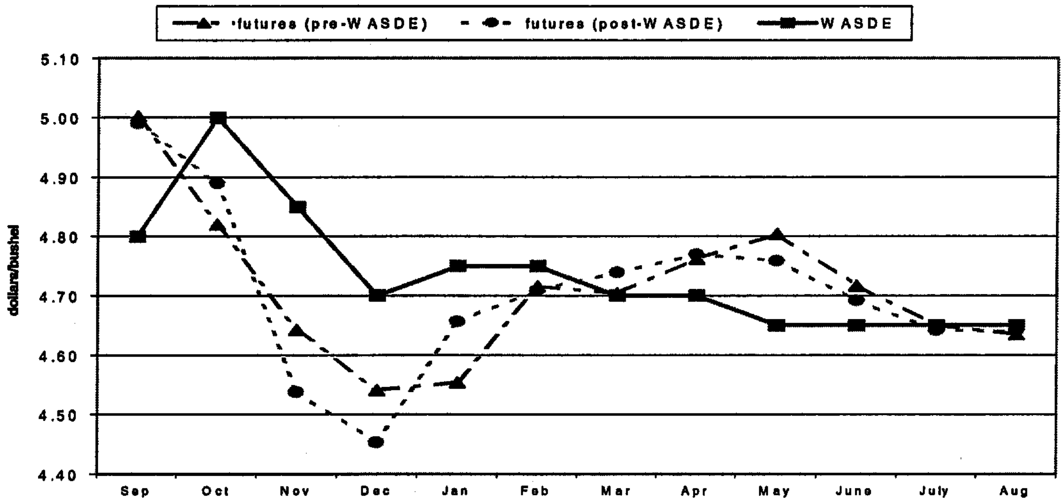
6. The NASS price estimates and forecast farm prices are spliced together in line 6. The November 1999 forecast of SAFP for crop year 1999/2000 will be based on actual price data for September and October, and forecasts for the remaining 10 months.

7. A five-year average of monthly marketing shares (in percents) by soybean producers (line 7) is used to weight the monthly farm prices (forecast or actual), yielding the final November 1999 forecast of the 1999/2000 soybean SAFP (line 8).

The November 1999 forecast of the 1999/2000 SAFP based on pre-WASDE futures information was $4.64/bushel. Although the actual 1999/2000 SAFP for soybeans is not yet available, this figure compares very favorably with the most recent (August 2000) WASDE point estimate of $4.65/bushel for the current crop year. In the months following the November forecast, the (pre-WASDE) futures forecast fell to about $4.55/bushel before climbing to a peak of just over $4.80 bushel in May 2000. The futures forecast then began to converge towards the WASDE estimate in June, July, and August (Figure 1).

The futures forecasts based on post-WASDE release futures data were all within about 10 cents per bushel of the pre-WASDE forecasts and the

Figure 1: Forecasts of U.S. Soybean Prices, 1999/2000 Crop Year



difference averaged about 4 cents/bushel. In November, the post-WASDE forecast was about 10 cents per bushel lower (at $4.54/bushel) than the pre-WASDE forecast. The difference is probably due to new information conveyed by the November WASDE report. USDA lowered its mid-point forecast of soybean SAFP by 15 cents per bushel due in part to diminished export prospects. The result was a less accurate forecast of the probable 1999/2000 soybean SAFP, but one still more accurate than the November WASDE mid-point projection of $4.85/bushel.

Compared to the WASDE price estimates, the futures price forecasts ranged from as much as 20 cents a bushel above the WASDE mid-point forecast in September 1999 to 31 cents a bushel below the WASDE projection in November 1999. Since the actual season average farm price for soybeans has not yet been established and just one year's worth of projections are represented here, these comparisons are somewhat less meaningful than the historical analysis of forecast accuracy for the crop years 1981/82 to 1998/99 presented in the next section.

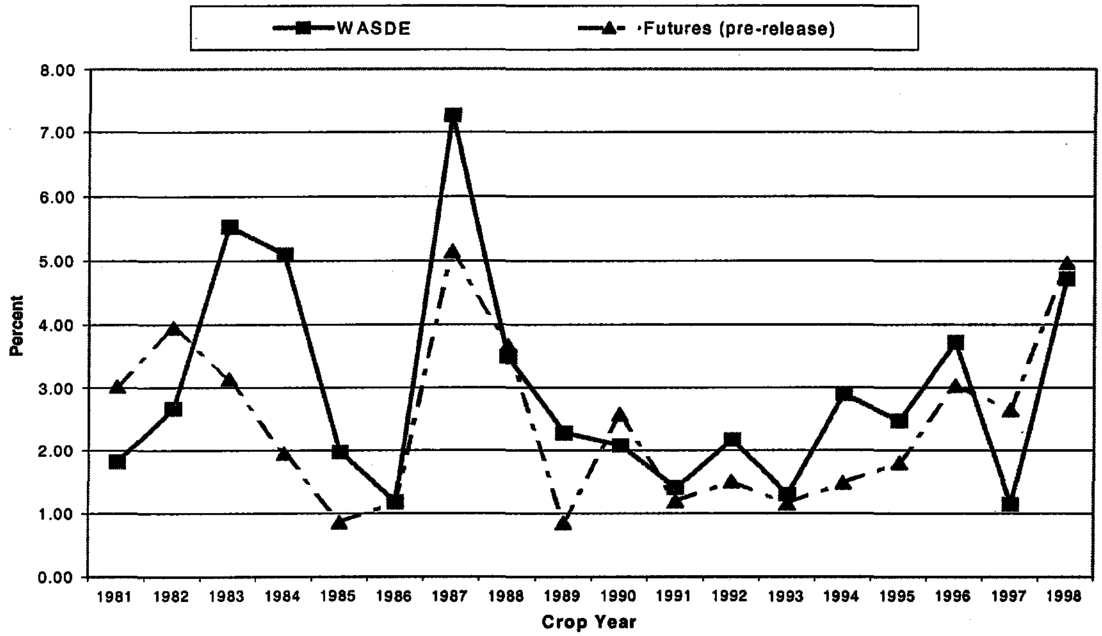
Forecast Accuracy of the futures method and WASDE (1981/82 to 1998/99)

In this section, we examine the historical (1981/82 to 1998/99) accuracy of soybean SAFP forecasts published in USDA's WASDE reports as well as the accuracy of the two alternative forecasts developed using the futures method. This analysis is designed to help us gauge the general accuracy of the WASDE projections, and to judge whether the futures method represents a reasonable alternative approach for developing such forecasts. Initially, forecast accuracy is assessed by calculating the mean absolute percentage error (MAPE) for each forecast (WASDE or futures) over the entire crop year. That is, for a given crop year, the MAPE gives the average percentage difference between each month's (September through August) forecast of SAFP and the actual SAFP. We then examine the average absolute percentage error of the monthly forecasts. For instance, the average absolute percentage error for the September WASDE report is the average of the September forecast errors over the 18 years examined. It should be remembered that the WASDE and futures forecasts of SAFP are composites of projected and actual (NASS estimates of) monthly cash prices as they become known.

Yearly forecast errors (1981/82 – 1998/99)

Figure 2 and the accompanying table present the mean absolute percentage errors for the WASDE and the futures method for crop years 1981/82 to 1998/99. The MAPE is a summary of monthly errors during each crop year and therefore masks fluctuations of the errors over the course of the crop year. Nevertheless, it provides a general sense of the overall accuracy of the alternative forecasts as well as a basis for comparison between the forecast methods. Since the results for the pre-WASDE and post-WASDE futures method were similar, figure 2 compares only the pre-WASDE futures forecasts

Figure 2: Mean Absolute Percentage Error (WASDE vs. Futures Method)



Mean Absolute Percentage Error (1981/82 – 1998/99)

Crop year	1981	1982	1983	1984	1985	1986	1987	1988	1989
WASDE	1.82	2.67	5.53	5.09	1.98	1.19	7.27	3.49	2.29
Futures (pre)	3.03	3.95	3.14	1.95	0.86	1.17	5.14	3.67	0.85
Futures (post)	2.95	3.88	3.06	1.51	0.56	1.25	4.80	3.27	0.68

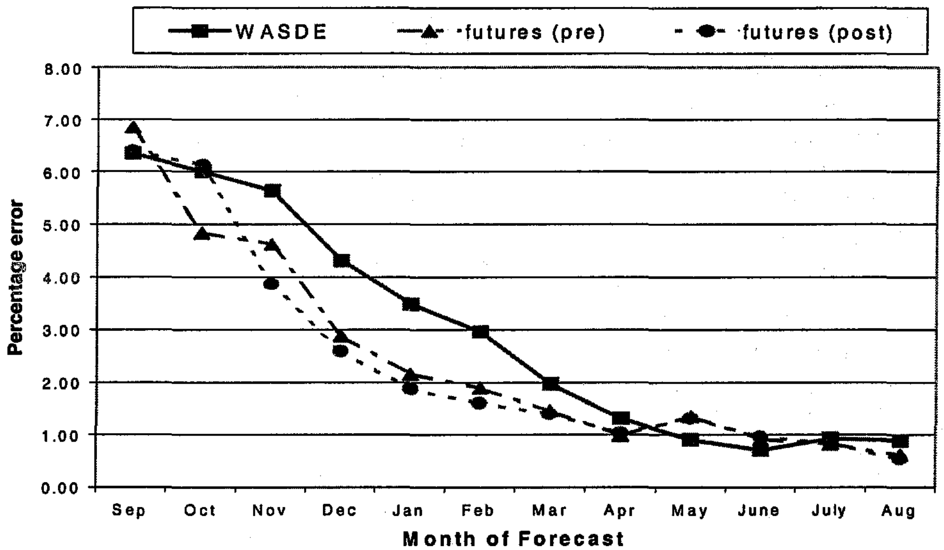
Crop year	1990	1991	1992	1993	1994	1995	1996	1997	1998	Mean
WASDE	2.09	1.40	2.17	1.30	2.90	2.48	3.72	1.15	4.72	2.96
Futures (pre)	2.58	1.19	1.51	1.16	1.50	1.80	3.04	2.65	4.97	2.45
Futures (post)	2.26	1.38	1.43	1.44	1.42	1.82	2.95	2.86	5.33	2.38

with the WASDE. The accompanying table provides the results for all three methods.

The MAPE for each of the three forecasts ranged from a low of 0.56 percent for the 1985/86 post-WASDE release futures method to a high of over 7 percent for the 1987/88 WASDE projections. By the MAPE criteria, it appears that the futures method holds a slight advantage over the WASDE in forecasting soybean SAFP. The average MAPE over the eighteen observations was 2.96 percent for the WASDE, 2.45 percent for the pre-WASDE release futures method, and 2.38 percent for the post-WASDE release futures method. The WASDE projection out-performed one or both futures forecasts in eight out of eighteen years, but in the other years, the WASDE errors tended to exceed those of the futures method by a fairly large margin – particularly in 1983, 1984, and 1987.

As indicated in Figure 2, the SAFP forecast errors for the WASDE and futures method tend to be highly correlated, generally falling or rising from previous year's errors in tandem. In addition, the tendency of all three forecasts was to somewhat overestimate soybean season average farm price. For each method, about 55 percent of the 216 monthly forecasts overestimated the final SAFP, but the simple mean error of all monthly forecasts was lowest for the WASDE (0.17 percent versus 0.36 percent for the pre-WASDE futures forecasts and 0.30 percent for the post-WASDE futures forecasts).

Figure 3: Average forecast error, by month of forecast (1981/82 - 1998/99)



Average absolute forecast error, by month of forecast (1981/82 – 1998/99)

Month	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
WASDE	6.35	6.01	5.65	4.31	3.49	2.96	1.98	1.31	0.91	0.71	0.94	0.89
Futures (pre)	6.85	4.84	4.62	2.89	2.16	1.90	1.46	0.99	1.35	0.92	0.83	0.62
Futures (post)	6.39	6.13	3.86	2.59	1.88	1.61	1.40	1.03	1.32	0.96	0.85	0.54

Monthly forecast errors (September – August)

Not surprisingly, the accuracy of SAFP forecasts for each method tends to improve over the course of the crop year, as actual monthly prices are incorporated into the forecasts. Interestingly, as shown in Figure 3 and accompanying table, the WASDE and futures method forecasts perform similarly during the first monthly projection (September) of the crop year SAFP. The eighteen-year average (of absolute) September forecast errors ranged from a low of 6.35 percent for the WASDE projection to a high of 6.85 percent for the pre-WASDE futures forecast. In the following months, particularly November through March, however, the WASDE projection errors consistently exceeded the futures forecast errors. Between November and February, the difference averaged more than 1 percentage point per month.

Why the WASDE forecast errors exceed the futures forecasts during these months is difficult to determine. One suggestion is that over the time period examined (1981/82 – 1998/99), WASDE projections of (U.S.) domestic use tended to be underestimated while ending stocks were overestimated. A look at statistics on the reliability of monthly WASDE projections between November and March (1981/82 to 1998/99) confirm this impression. The expected impact would be a consistent underestimation of the SAFP, but a closer look at monthly WASDE forecast errors does not support this conclusion. The simple average of errors for November, December, and January were positive, meaning price forecasts were slightly overestimated during these months. In any event, this suggestion does not explain differing magnitudes of WASDE and futures method forecast errors, only a potential pattern to WASDE forecast errors (which is not apparent).

Another suggestion is that the difference between WASDE and futures method forecast errors from November to February may be related to uncertainties about South American soybean production. Soybean planting in South America typically occurs in October, with harvest beginning in March. Less accurate or timely information on these crops could contribute to forecasting errors, but again, it is unclear that this would have a greater impact on WASDE forecasts than those based on the futures method.

It should be pointed out that, regardless of the source of the WASDE forecast errors, the accuracy of WASDE forecasts made during November through March have improved significantly during the 1990s, while those of the futures method have actually worsened slightly. Compared to the 1980s (1981/82-1989/90) time period, the average November-March WASDE forecast error decreased by more than 1 percentage point in the 1990s (1990/91-1998/99), whereas futures forecast errors increased by a little more than 0.1 percentage points during the same interval. This may reflect improved information, analysis, or modeling efforts by the USDA.

Summary and Conclusion

The goals of this analysis were twofold: to develop and illustrate the use of the futures method model for forecasting season-average farm price for soybeans, and to assess and compare the historical accuracy of this method with USDA's farm price forecasts published monthly in WASDE. Our findings suggest that both the WASDE and futures method provide reasonable and generally accurate price forecasts. By the mean absolute percentage error (MAPE) criteria, the futures method slightly outperformed the WASDE projections, but a simple average of all (216) monthly forecast errors indicates that the WASDE does not overestimate the SAFP as much as futures method forecasts. In addition, there is little to distinguish the WASDE from the futures method in terms of beginning-of-the-crop-year accuracy. The futures method is typically more accurate between November and March of the crop year, but the differences are narrowing. Finally, the MAPE of futures forecasts based on post-WASDE release futures prices are on average lower than pre-WASDE futures forecasts – indicating that information conveyed by WASDE reports improve futures method forecasts – but the difference is minor.

In conclusion, the futures method of forecasting the season-average-farm-price of agricultural commodities represents a useful tool for analysts and market participants seeking a cross-check to USDA projections. Future research on the method could examine alternative methods of estimating the basis and marketing weights, such as using a five-year moving olympic average (omitting the high and low figures) rather than a simple moving average. Improved estimates of these variables should enhance the overall accuracy of price forecasts. Another avenue would be to examine the historical accuracy of other forecasting tools that have been used to project commodity prices, such as time series (autoregressive-integrated-moving-average) models. Using the ARIMA method, Vroomen and Douvelis (1993) developed forecasts of soybean SAFP for crop years 1989/90 to 1991/92 with results similar to WASDE and futures method forecasts, but it is unclear whether the accuracy of this method would be sustained over the longer run.

References

Hoffman, L, and C. Davison. "Forecasting U.S. Soybean Prices with Futures Prices." *Oil Crops: Situation and Outlook Report,* ERS, USDA. January, 1992.

Hoffman, L. and J. Balagtas. "Providing Timely Farm Price Forecasts: Using Wheat Futures Prices To Forecast U.S. Wheat Prices At the Farm Level." *The 10th Federal Forecasters Conference – 1999: Papers and Proceedings.* Bureau of Labor Statistics, Washington, DC, June 24, 1999.

U.S. Department of Agriculture, National Agricultural Statistics Service. *Agricultural Prices.* Various issues.

U.S. Department of Agriculture, National Agricultural Statistics Service. *Crop Production.* Various issues.

U.S. Department of Agriculture, World Agricultural Outlook Board. *World Agricultural Supply and Demand Estimates.* Monthly issues, 1976-2000.

Vroomen, H. and G. Douvelis. "A Short-Run Forecasting Model for Soybean Prices" *Oil Crops: Situation and Outlook Report,* ERS, USDA. January, 1993.

COINTEGRATION TESTS AND PRICE LINKAGES IN WORLD COTTON MARKETS

Stephen MacDonald, Economic Research Service, USDA

Abstract

Cotton is a tradable good, and the volume of U.S. trade suggests significant incentives for the integration of U.S. and world cotton markets. During the 1980's, the U.S. share of world cotton trade averaged 21 percent and the exported share of U.S. cotton production averaged 48 percent. During the 1990's, the respective shares were 25 and 40 percent. However, cointegration analysis of the relationship between U.S. and world prices finds varying evidence for the integration of U.S. and world cotton markets, with the law of one price apparently violated during the 1990's.

Introduction

Before 1985, the U.S. farm policy acted to segregate U.S. farm prices from world prices through high loan rates. Thus, during this period, the accumulation of U.S. government-owned stocks served to prevent the transmission of price signals between U.S. and world markets. Bessler and Chen, testing the relationship between the monthly A-Index (world cotton price, Northern Europe) and the monthly Memphis cotton quote in Northern Europe, found the prices were cointegrated during January 1980-November 1994, but note that, "Whatever long-run relationship that did exist in the pre-1985 data, it was not particularly strong."

With the implementation of the marketing loan program, U.S. prices were free to adjust to below the loan rate, an important step in market integration. With the 1990 U.S. farm legislation, a mechanism for expanded U.S. import quotas was created, further increasing the opportunities for arbitrage between U.S. and world markets. Events during the 1990's have demonstrated that the market access provided by the special import quotas is real. With 80 consecutive weeks of special import quotas opening through May 1997, U.S. cotton imports reached amounts unmatched in 70 years (MacDonald). During March-December 1996, imports totaled more than 700,000 bales, compared with 1,000 to 20,000 bales per year during the preceding decade. In marketing year 1998/99 the United States imported 443,000 bales. Since 1995, imports have accounted for 1.8 percent of U.S. cotton consumption, compared with 0.1 percent during the decade preceding.

However, the 1990 legislation also created User Marketing Certificates for U.S. cotton (a program generally referred to as "Step 2"). Step 2 results in payments to U.S. mills and exporters using U.S. cotton during periods when U.S. prices exceed world prices, when certain conditions hold. The magnitude of the payments is determined by the magnitude of the difference between U.S. and world prices. (see Glade, Meyer, and MacDonald for background). During the 1990's, Step 2 payments averaged $199 million per year, ranging from $3 million to $422 million (USDA, Farm Service Agency). Step 2 payments might be expected to weaken the integration between U.S. and world cotton markets.

During the 1990's there was also an important change in world markets, with the emergence of Central Asia as the largest competitor for the United States. Before 1990, Central Asia's cotton was largely consumed within the COMECON countries of Eastern Europe and the Soviet Union, and had limited impact on cotton trading in the rest of the world. With the economic reorientation of these countries, and the collapse of Russia's textile industry, a new, low-cost competitor of enormous proportions appeared on world cotton markets. According to the International Monetary Fund, the governments of major cotton producers in Central Asia—Uzbekistan and Turkmenistan—acquire virtually the entire local cotton crop at well below world prices through either state orders or export controls. Thus, Central Asian cotton is typically the least expensive cotton available on world markets, and, with a 25 percent share of world trade during the 1990's, clearly exerts an important influence. Over the last decade, the accumulated impact of environmental damage and autarkic economic policies has in part resulted in a steady decline in the region's output and exports, adding a dynamic factor to its influence on world markets.

In this paper, U.S. and world cotton prices are examined for stationarity and cointegration, and evidence of structural change since 1991 and the violation of the law of one price since then is presented.

Previous Research

Bessler and Chen do not report their findings concerning price stationarity in their study covering 1977-93. Baffes and Ajwad report mixed results using the standard stationarity tests. They find consistent evidence of non-stationarity for 1985-87, but their tests over 1995-97 show trend stationarity, but non-stationarity when a time trend is excluded. Baffes and Ajwad also apply a variance-ratio test the results of which point to non-stationarity, and report that the cumulative evidence supports the conclusion of non-stationarity.

Bessler and Chen found U.S. and world cotton prices were cointegrated both during 1980-1984 and 1986-1993. They noted an interval during 1985-86 where cointegration was evidently not operating, an interregum that is readily observable in a graph of the difference between the A-Index and the U.S. spot price (Figure 1). The disruption caused by the transition from U.S. price supports through loan rates to the marketing loan program in place since 1986 affected the relationship between world and U.S. prices.

While Bessler and Chen find U.S. and world prices are cointegrated during both time periods, they qualify their results for 1977-84, citing a failure to reject weak exogeneity for both series during that time period.

Baffes and Ajwad do not directly report results for cointegration, instead analyzing "comovement" given an assumed cointegration parameter. No comovement was reported between the A-Index and the Memphis price over August 1985—December 1987, using weekly observations, but a high degree of comovement was observed during August 1995—January 1997. Similarly, they estimate error-correction models for these two time periods, and observe no long-run relationship between the A-Index and the Memphis price in the first period and the presence of a long-run relationship for the later period.

Thus, both studies support the conclusion that U.S. and world prices were not linked during 1985/86 and were linked during 1986-1991. Both studies used the Northern European (N.E.) quotes for Memphis cotton for their U.S. price.

Data

Prices examined in this paper are the monthly August 1986—December 1999 U.S. average spot price published by USDA's Agricultural Marketing Service (AMS) and the A-Index of Northern European quotes, published by Cotlook Ltd. Complete descriptions of each price series can be found in Larson and Meyer, which are summarized below.

The average spot market price is the average quoted for the base quality in each of seven U.S. marketing areas. AMS cotton market reporters gather market news in person and by telephone, and in the absence of trading in a particular market, quotations are determined by prices paid for similar qualities in other markets. Because spot prices are simple averages they may be skewed by aberrant prices in markets with low trading volumes. The base staple-length of the spot price is 1 1/16th inches.

The Cotlook A-Index® is based on a Liverpool concept of Middling 1-3/32 inch staple-length cotton. At the close of trading each day, Cotlook Ltd.'s Memphis office collects offering prices across the United States from merchants who trade in the international market. The Liverpool office collects similar prices in Europe, and a market value of various descriptions of cotton is determined daily from this information (e.g., for U.S. Memphis, U.S. California, Chinese Type 329, Pakistani Punjab SG 1503. See *Cotton Outlook* for a current complete list.). The average of the 5 lowest-priced descriptions out of a basket of 15 comprise the A-Index. The A-Index is not weighted by quantity traded, and shipment dates can often vary by months between descriptions. Since the A-Index is not comprised of a fixed basket of prices, it can vary as reduced availability terminates quotations for a certain description of cotton for the year. This can result in large day-to-day shifts in the A-Index level as the unavailability of quotes in the lowest price cotton will result in the substitution of a high-priced growth in the average.

The A-Index quotations are also specific to the fiber's year of production. This introduces a discontinuity into the price series used here since the A-Index for a given July refers to cotton produced in marketing year X and the subsequent price for August refers to year X+1. A forward A-Index referring to the coming marketing year (X+1) is available during the latter part of each marketing year X, but in this study no adjustment based on these forward quotes was used.

Results

Tests for the presence of unit roots and cointegration are now commonly elucidated in econometric texts and incorporated into statistical software, so the basis and nature of these tests will not be elaborated upon here. See Harris for an introduction, and Banerjee, et al, for a more complete exegesis.

Both the augmented Dicky-Fuller (ADF) and Philips-Perron (PP) tests indicate that both prices are I(1) in virtually every case (Tables 1 and 2). The A-Index during 1986-91 is the only exception, with the ADF and PP results suggesting, respectively, rejecting the null hypothesis of non-stationarity and accepting it. Examining the sample's data (Figure 2) suggests the A-Index might appear to follow a trend during the period analyzed, and both the ADF and PP tests support the conclusion that the A-Index is trend stationary during 1986-91. Given that most price series tend to be non-stationary, and that the evidence is mixed in this case, the analysis proceeded under the assumption of non-stationarity even without a trend. The cointegration results are the same in each case.

No formal tests were made for the timing of a structural break. The first marketing year under 1990 farm legislation marked an important shift in the policy regime of the world's largest exporter, suggesting an appropriate break point. Figure 1 also suggests change in the relationship between U.S. and world prices at about that time. After 1991 the gap between U.S. and world prices narrows.

The A-Index averages quotes for cotton 1/32 inch longer than that priced by the U.S. spot price, suggesting a premium for the A-Index based on quality. Northern Europe is also relatively distant from regions of significant cotton production, and the cost of transportation between the United States and Northern Europe would be expected to add a further premium to the A-Index. Transportation costs are calculated annually by USDA's Economic Research Service and are generally nearly 14 cents per pound.

Note that the A-Index's premium is seldom large enough to encompass both of these factors. U.S. cotton of a given specification and location is generally higher-priced than that of another country due to reliability and quality factors.

Comparison between Figures 2 & 3 illustrate the differences between the two time periods. During the earlier period, the A-Index exceeded the spot price in every month, with the exception of May 1991. The average premium for the A-Index was 9.5 cent per pound. During 1991-99 the average premium fell to 3.8 cents, with the spot price actually exceeding the A-Index for an extended period in 1998.

If the prices were cointegrated during each period, the differences could be attributed to a change in the intercept of the cointegrating relationship. Perhaps the greater role of Central Asian cotton in determining the A-Index increased the U.S. premium relative to the rest of the world on average. Similarly, payments under the Step 2 program could shift the premium between cotton on U.S. and world markets.

However, cointegration testing indicates that U.S. and world prices after 1991 are no longer cointegrated. Rather than just altering the difference between the average levels of the two prices, changes in world markets have altered the relationship between the two prices. Over the entire period studied (1986-99), cointegration appears to hold (Table 3). Similarly, during the earlier period, 1986-91, the prices appear to be cointegrated (Table 4). However, since 1991, the null hypothesis that there are no co-integrating vectors cannot be rejected (Table 5). These results are all robust across a range of vector-autoregression (VAR) model lags and specifications with respect to intercepts and trends in both the VAR models and the cointegrating equations. Log-likelihood ratio tests indicated VAR lags for the three respective time periods of 12, 9, and 2 months.

Conclusions

The relationship between U.S. spot prices and the A-Index seems to have changed since 1991, although further research will be necessary to determine the sources of this change. The differences in the average price gap, VAR lag length, and ADF lag lengths all point to possible structural change between 1986-91 and 1991-99. The disappearance during the latter period of a cointegrating relationship observed during the earlier period supports this conclusion. Interestingly enough, the change in the relationship is not sufficient to result in an apparent lack of a non-cointegrating relationship when estimating over the entire 1986-99 time period.

The variability in the relationship between the two prices during 1991-99 could have several sources. The Step 2 program, for example, could be understood to affect the relationship in two ways. One way is by sundering the link between U.S. and world markets. If User Marketing Certificates are typically available to equate U.S. and world prices for cotton exporters and consumers when these prices diverge, then the pressure of arbitrage to bring them together again is lessened.

These certificates are not always available, but variability in the operation of Step 2 may have led to changes in the U.S./world price relationship within the 1991-99 period. Initially, Step 2 payments for exports were based on the prevailing certificate value on the date of sale. In 1995, a regulatory change shifted the export payment determination from date of sale to date of shipment. Under the original regulations, millions of bales were sold for export during a single week in response to a perceived peak in certificate values. Most Step 2 payments went to exporters during that period.

The 1996 U.S. farm legislation added a cap to Step 2 expenditures of $701 million through 2002. Previously, potential expenditure was unlimited. The expenditure limit was reached in December 1998, and efforts to consummate shipments before the exhaustion of funds introduced some unusual price dynamics during marketing year 1998/99. Legislation in 1999 removed the spending cap, and the relationship between Step 2 payments and Special Import Quotas ("Step 3") was adjusted.

This summary of major changes in U.S. cotton policy indicates that even if policies like Step 2 did not break the link between U.S. and world prices, the nature of that link could have changed several times during 1991-99, resulting in an apparent absence of cointegration due to structural breaks.

Another factor which may have introduced instability into the U.S./world price relationship has been the varying role of Central Asia in world cotton markets. Early in the 1990's, barter arrangements developed before the collapse of the Soviet Union accounted for a substantial portion of Uzbekistan's and Turkmenistan's exports. The last of these agreements only lapsed late in the 1990's. Price transmission between Central Asia and the rest of the world would probably vary as the role of barter varied. Furthermore, as noted in this paper's introduction, Central Asian production and exports have been declining over the 1990's, varying the region's impact on the A-Index and other prices.

It may be that the A-Index is an inappropriate variable for the tests used in this paper. Rather than a fluctuating basket of prices, it may be appropriate to test for the integration of a specific foreign price with the U.S. price. On the other hand, the A-Index is widely recognized in the industry as the world price, and is identical to the price index used in determining the value of Step 2 certificate values and the use of Special Import Quotas. This suggests that further research involving the A-Index would be at least as useful as that with substitute prices.

References:

Baffes, John, and Ajwad, Mohamed I., "*Detecting Prices Linkages: Methodological Issues and An Application to the World Market of Cotton*," World Bank, Research Policy Working Papers. July 1998.

Banerjee, Anindya, Dolado, Juan, Galbraith, Jown W., and Hendry, Davide F., *Co-integration, Error-Correction, and the Econometric Analysis of Non-Stationary Data*, Oxford University Press. 1997.

Cotton Outlook, Cotlook Ltd., Liverpool, U.K.

Bessler, David A., and Chen, Dean T., "Long-Run and Short-Run Relationships Between 'A-Index' and Memphis Prices," *Proceedings of the 1995 Cotton Beltwide Conferences*. National Cotton Council. 1995.

Glade, Edward H., Meyer, Leslie A., and MacDonald, Stephen, *Cotton: Background for 1995 Farm Legislation*. USDA, ERS. 1995.

International Monetary Fund, *Republic of Uzbekistan: Recent Economic Developments*, IMF Staff Country Report No. 00/36. 2000.

International Monetary Fund, T*urkmenistan: Recent Economic Developments*, IMF Staff Country Report No. 99/140. 1999.

Johansen, Soren, and Katerina Juselius, "Maximum Likelihood Estimation and Inference on Cointegration with Application to the Demand for Money," *Oxford Bulletin of Economics and Statistics* 52: 169-209. 1990.

Larson, James A. and Meyer, Leslie A., "Supply, Demand, and Prices," In , Glade, Edward H., Meyer, Leslie A., and Stultz, Harold (ed.*) The Cotton Industry in the United States*, USDA, ERS. 1996.

Harris, Richard I D., *Using Cointegration Analysis in Econometric Modeling*, Prentice Hall. 1995.

MacDonald, Stephen, "Cotton Marketing Loan Program Shifts Away from Imports," *Agricultural Outlook*, AO-242, USDA, ERS. July 1997.

Osterwald-Lenum, Michael, "A Note with Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Test Statistics," *Oxford Bulletin of Statistics* 54: 461-71. 1992.

USDA, Farm Service Agency, "Fact Sheet, Upland Cotton: Summary of 1999 Commodity Loan and Payment Program." April 2001.

Table 1—U.S. Spot Prices, Unit Root Tests

	Lag	ADF	PP	ADF	PP
	Levels			1st Diff.	
1986-99	1	2.64	2.54	8.70**	10.72**
1986-91	0	2.26	2.42	7.28**	7.26**
1991-99	4	1.57	1.49	5.21**	6.92**

** significant at 1%
(Lag refers to lag of the preferred ADF model. PP lags were determined by Newey-West automatic truncation selection.)

Table 2—A-Index, Unit Root Tests

	Lag	ADF	PP	ADF	PP
	Levels			1st Diff.	
1986-99	4	2.65	2.70	6.21**	7.69**
1986-91	1	3.25*	2.84	5.16**	7.20**
1991-99	4	1.26	1.25	4.93**	6.12**

*significant at 5%, ** significant at 1%
(Lag refers to lag of the preferred ADF model. PP lags were determined by Newey-West automatic truncation selection.)

Table 3—Johansen Cointegration Test Summary, 1986-99.

Eigen-value	Likelihood Ratio	5% Critical Value	r
0.255	44.049**	24.60	0
0.007	1.029	12.97	1

* *rejection at 1% significance of null hypothesis that largest number of cointegrating relationships = **r**

Table 4—Johansen Cointegration Test Summary, 1986-91.

Eigen-value	Likelihood Ratio	5% Critical Value	r
0.763	89.332**	19.96	0
0.057	2.878	9.24	1

* *rejection at 1% significance of null hypothesis that largest number of cointegrating relationships = **r**

Table 5—Johansen Cointegration Test Summary, 1991-99.

Eigen-value	Likelihood Ratio	5% Critical Value	r
0.067	7.616	19.96	0
0.010	0.972	9.24	1

* rejection at 5% significance of null hypothesis that largest number of cointegrating relationships = **r**

Figure 1: A-Index Price Premium 1976-2000

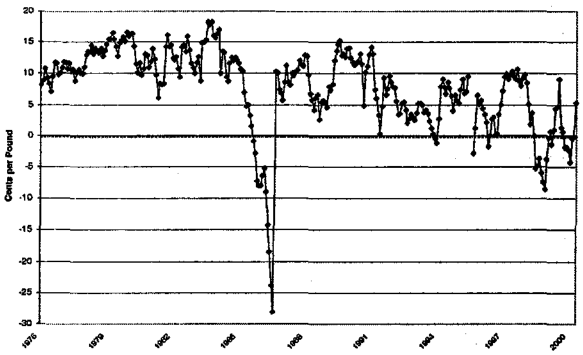


Figure 2: U.S. and World Cotton Prices 1986-91

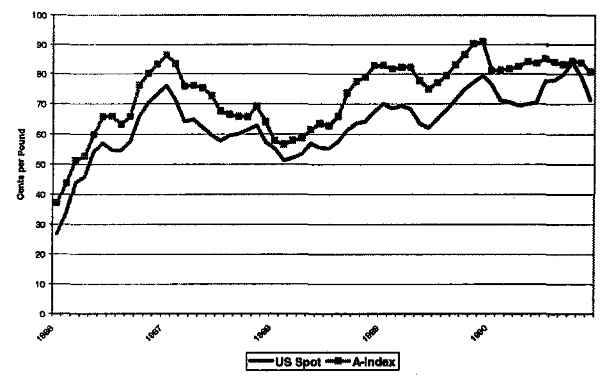
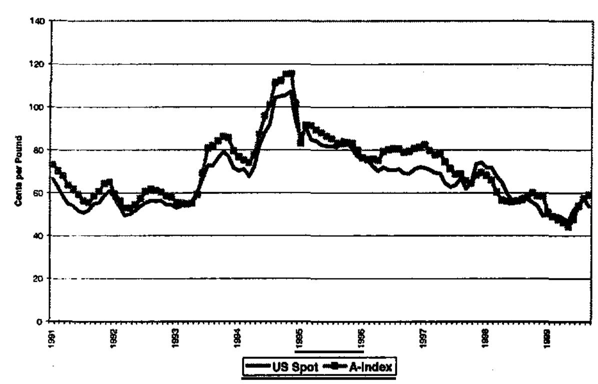


Figure 3: U.S. and World Cotton Prices 1991-99



☆U.S. GOVERNMENT PRINTING OFFICE: 2001-480-384/40227